PA 1091313

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

November 10, 2003

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE UNDER 35 USC 111.

APPLICATION NUMBER: *60/429,532*
FILING DATE: *November 29, 2002*

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

P. SWAIN
Certifying Officer

Approved for use through 01/31/98. OMB 0651-0037
Patent and Trademark Office; US DEPARTMENT OF COMMERCE

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

A/PROV

This is a request for filing a **PROVISIONAL APPLICATION FOR PATENT** under 37 CFR 1.53 (b)(2).

Docket Number	25102	Type a plus sign (+) inside this box ->	+

INVENTOR(s) / APPLICANT(s)

LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)
SMILANSKY	ZEEV		Yavne, Israel

TITLE OF THE INVENTION (280 characters max)

PROTEIN SYNTHESIS MONITORING (PSM)

CORRESPONDENCE ADDRESS

G. E. EHRLICH (1995) LTD.
c/o ANTHONY CASTORINA
2001 JEFFERSON DAVIS HIGHWAY
SUITE 207

STATE	VIRGINIA	ZIP CODE	22202	COUNTRY	USA

ENCLOSED APPLICATION PARTS *(check all that apply)*

☑ Specification	*Number of Pages*	50	☑ Applicant is entitled to Small Entity Status	
☑ Drawing(s)	*Number of Sheets*	12	☐ Other (specify)	

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT *(check one)*

☐ A check or money order is enclosed to cover the filing fees		**FILING FEE AMOUNT ($)**	**$ 80.-**
☑ The Commissioner is hereby authorized to charge filing fees and credit Deposit Account Number:	**50-1407**		

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☑ No

☐ Yes, the name of the US Government agency and the Government contract number are: _____

Respectfully submitted,

SIGNATURE _____ Date 26 November 2002

TYPED or PRINTED NAME SOL SHEINBEIN

REGISTRATION NO. (if appropriate) 25,457

☐ Additional inventors are being named on separately numbered sheets attached hereto

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

Burden House Statement: This form is estimated to take 2 hours to complete. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington DC 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, DC 20231.

November 26, 2002

Inventor: Zeev Smilansky

PROTEIN SYNTHESIS MONITORING (PSM)

FIELD

This disclosure relates to the monitoring of protein synthesis by ribosomes and may be used in biology and pharmacology.

BACKGROUND

Introduction

The study of proteins is quickly becoming a key endeavor of current biological research, as well as a focus of pharmaceutical research and development. The information revealed by sequenced genomes serves mainly to fuel protein research, be it the development of a cell-based assay, analysis of a pathway, study of a single receptor, or application of proteomics (the study of the entire protein complement of a tissue or organism). Current technology fails at several key points: it is expensive, difficult and slow; misses entire protein families; fails to identify protein pathways; focuses on a single protein at a time; and, importantly, provides very little information on protein dynamics. In fact, results of current large-scale and high-throughput protein analysis are often delayed by days or weeks following an experiment, and are usually restricted to the form of a catalogue, tabulating those proteins of a database that have been putatively identified from the analyzed sample.

Genomics, Proteomics and the Barriers of Biological Knowledge

Proteomics is an emerging technology that attempts to study proteins on a large scale in high-throughput. It is not by chance that the term resembles

"Genomics". In the wake of successful technologies such as whole genome sequencing, DNA chips and SNP cataloging, people began looking for similar paradigms in the realm of proteins. This is worthwhile since proteins are so much more important than DNA, being the main vehicles of life processes: they are the biochemical enzymes, form the signal pathways, control the cellular processes, underpin the cell scaffolding, transport molecules and so on. They are also much more valuable than DNA, due to their importance in human disease: most known drugs are either proteins themselves, or else operate by binding to a protein target. Unfortunately, proteins are also so much more complex and difficult to study than DNA. They are more complex since there are a lot more proteins than there are genes; since protein expression is complex and has a high dynamic range – from single copies to millions per cell; since the proteome of one cell type may be very different than that of another, even though their DNA is identical; and since proteins lead a rich and eventful life, during which their structure may change dramatically – through cleavage, modification, and interaction. Proteins are more difficult to study than DNA, since protein extraction, separation and identification are difficult; since there is no amplification technique that parallels PCR; since protein 3d-structure is hard to obtain and use; since protein expression has high dynamic range; since protein modifications, cleavages and interactions are to a large extent unknown; and, finally, both as a result and a reason, since protein databases are thin, encompassing a small fraction of all theoretical proteins, especially in higher organisms, such as *homo sapiens*.

In one aspect, though, Proteomics and Genomics proved similar: both raised high hopes of creating a paradigm shift, a breakthrough that will yield a new understanding of cellular processes and human disease, and pave the way to a bounty of new drugs and therapeutics. Unfortunately, first for genomics and then for proteomics, it became abundantly clear that though genomic and proteomic data is extremely valuable, it is far from sufficient for achieving the breakthrough that was hoped for (Miklos, G.L. and Maleszka, R., *Protein functions and biological contexts*. Electrophoresis 22:169-178, 2001). So many

puzzle pieces are still missing that the clear view of cellular machinery remains hidden from us. One important piece of this puzzle is protein synthesis data- which proteins are produced at what times, under which conditions, and in what amounts. The ability to study and monitor this type of data will be a major breakthrough for all life science related research – though this may sound less like science and more like science fiction to researchers today.

Proteomics Practice Today

Mainstream proteomic analysis today goes through the steps of protein purification from culture, separation with 2D-gel or other chromatographic techniques, mass-spectrometry, and analysis of the resulting spectra for protein identification and characterization.

The extraction of proteins from bacterial or cell culture invariably involves lysis (and therefore death) of the cells. The procedure involves several steps and usually takes hours (Branca MA, Sannes LJ. *Proteomics: A Key Enabling Tool for Genomics?* Cambridge Healthtech Institute's Genomic Reports. April 1999; Humphery-Smith I., Cordwell S.J., Blackstock W.P., *Proteome research: complementarity and limitations with respect to the RNA and DNA worlds*, Electrophoresis 18 (1997) 1217-1242). Protein separation with 2D-Gels requires at least 24 hours and an expert; their analysis is often much more difficult, even with modern software (Smilansky, Z. *Automatic registration for images of two-dimensional protein gels*, Electrophoresis 2001, 22, 1616–1626). Even worse, 2D-Gel technology is not applicable to very acidic or very basic proteins, to many membranal proteins, and most importantly, to proteins that are expressed in low amounts.

It is usually taken for granted that proteins that are expressed at less than 1000-10,000 copies per cell cannot be visualized in 2D-Gels (Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., Aebersold, R., *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags*. Nat Biotechnol. 1999, 17(10): 994-9). Almost no protein kinases, phosphatases, transcription factors, GPCRs, ion channels, or nuclear hormone receptors are found in standard

human proteomic analyses, even though more than 5000 of these proteins are encoded by the human genome (Miklos, G.L., Maleszka, R. *Protein functions and biological contexts*, Electrophoresis 22:169-178, 2001). Thus, the proteins that can be analyzed by this method are the most common, and therefore, often the least interesting ones.

Besides separating the sample, 2D-Gel technology can measure three important protein parameters: mass, pI, and quantity. However, all three are hopelessly inaccurate. As for protein quantity, the most one can hope for from gel technology is relative quantitation, and even that at accuracies worse than 50% error – so that only proteins with very strong up- or down-regulation can be identified. Moreover, quantitation at best means quantity of protein in the extracted, processed sample, such as in a gel spot or in a chromatographic fraction; Estimation of protein copies in a cell at any given time is not even attempted today.

Following protein separation, MS analysis usually takes place, either with a MALDI-TOF or with an LC-MS-MS machine (Humphery-Smith I., Cordwell S.J., Blackstock W.P., *Proteome research: complementarity and limitations with respect to the RNA and DNA worlds*, Electrophoresis 18 (1997) 1217-1242; Yates J.R., *Database searching using mass spectrometry data.* Electrophoresis 1998, 19(6): 893-900). The main steps are spot picking from the gel followed by destaining, or alternatively chromatographic prefractionation, followed by protein digestion with a protease (almost invariably trypsin), mass-spectrometric analysis, and finally database searching, which is carried out, amazingly enough, only as a semi-automatic procedure with expert supervision and decision making – as in the steps of peak extraction and candidate selection.

All in all, the standard technique for identifying proteins in a cell culture takes from weeks to months, is suitable for only a small part of the proteome, does a bad job of quantitating protein amounts, and provides no clue as to the proteome dynamics.

Additional and Emerging Proteomics Technologies

An important older method for protein analysis is Edman degradation, a chemical analysis method where the C-terminal amino acids of a polypeptide are cleaved and analyzed one by one. The procedure requires a full day and provides no quantitative or dynamic information.

The problems of 2D-Gel technology have led many researchers to look for alternatives. Two important developments of the last few years are the techniques of ICAT (Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H., Aebersold, R., *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags.* Nat Biotechnol. 1999, 17(10): 994-9) and MudPIT (Washburn, M.P., Wolters, D., Yates JR 3rd. *Large-scale analysis of the yeast proteome by multidimensional protein identification technology.* Nat Biotechnol. 2001 Mar 19(3): 242-7), which involve MS analysis of whole sample digestion products. The two methods allow better identification of rare proteins, and the first one even allows computation of differential expression. However, they are still difficult and expensive to carry out, require cell lysis, take days for complete analysis, and provide no dynamic information.

Protein chips are being developed in several labs (Jenkins, R.E. and Pennington, S.R. *Arrays for protein expression profiling: towards a viable alternative to two-dimensional gel electrophoresis?* Proteomics. 2001 Jan 1(1):13-29). They generally fall into one of three classes: surface chemistry chips, antibody chips, or protein chips for determining protein-protein interactions. All of these may aid protein analysis in some way, but none of them provides the data that the disclosed method provides.

Yeast-two-hybrid technique (Y2H) is a feat of bioengineering that helps discover protein-protein interactions (Legrain, P. and Selig, L., *Genome-wide protein interaction maps using two-hybrid systems.* FEBS Lett. 2000 Aug 25; 480(1):32-6). The method is indirect in that the interactions occur in yeast or in bacteria, rather than in the original cells being analyzed. It is known to generate a large number of false-positives. The indirectness, together with lack of dynamic information, places it widely apart from the disclosed method.

High Throughput Screening and Cell Based Assays

High throughput Screening (HTS) is becoming the standard route for drug discovery in the pharmaceutical industry. Traditionally, HTS relies on a simple assay, such as receptor binding or enzyme activity. The assay itself measures receptor binding, that is, the efficacy of the candidate compound as an agonist or antagonist to the target protein. The rest of the required information – ADME-TOX for example – is either presumed to be known or its acquisition is delayed till later stages in the process (see also next section).

Cell-based assays have been in use for many years in the pharmaceutical industry. They are usually used for lead optimization and predictive toxicology. To construct a cell-based assay, a measurable characteristic has to be developed: this can be a fluorescent-tagged protein, an antibody based marker, or a measurable phenotypic characteristic of the cell. Modern examples include cancer-specific dyes (http://www.zetiq.com/site/cama.html) and genetically engineered cell lines (Shen-Orr, S.S., Milo, R., Mangan, S., and Alon, U., *Network motifs in the transcriptional regulation network of Escherichia coli.* Nat Genet 2002, 31(1): 64-8; http://www.cellomics.com/).

Cell based assays have many advantages over receptor binding assays. Cells offer better representations of a disease. By screening against disease pathways in whole cells, no a-priori assumptions are made about what makes a good target. However, cell based screening suffers from the need to engineer a specific cell line with the required reporting capability. This once again limits the scope of information obtainable. And of course, the entire cascade of events following administration of the compound under analysis remains hidden from the researcher. This is in strong contrast to the assay method disclosed here, where upon administration of the compound no special preparation are required (besides using the cells prepared in a standard way), no assumptions whatsoever are made, and protein synthesis processes can be followed to gain a complete understanding of the cell's response to the chemical or environmental stimulus that was applied.

The Need for the present invention

It is agreed that pharmaceutical R&D is today in crisis. Despite the huge increase in investments and the enormous contributions of genomics and related technologies, the main difference between the pharmaceutical pipeline today and a decade ago is in the number of targets, while the number of successful drugs entering the market has more or less stayed the same. More discouraging yet is the fact that while advances in high-throughput screening, chemical compound library design and bioinformatics have helped multiply the number of "hits" in HTS assays, the number of "leads" has not increased at all. Thus, the pharmaceutical pipeline today has an abundance of targets on the one side and an abundance of candidate compounds on the other, but the head-on collision of these two armadas has yielded little.

Though there is more than one reason for this, one important point is that though the numbers of targets and candidates is huge, the complexity of the cellular machinery, not to mention tissue and whole organism, is on a grander scale still. Thus, what is missing is a better view of function and context of a protein target in the cell, as well as the complex effects, side effects, and after effects of a drug compound on the cell.

In today's paradigm of drug development, once a target is found and a compound that binds to it identified, the huge drug development machinery starts rolling on the long road towards regulatory approval and market acceptance. While the road is long and very expensive, it is narrow in the sense that not a lot is known about the target protein, its function, its isoforms and look-alikes, its roles in disease and in health; Even less is known about the drug candidate, how it effects proteins other then its specified target, how it effects other tissues, what are its immediate effects and what are its later effects. Thus, it often turns out that information that can wipe a compound off the drug list is revealed only at later stages and at a high cost – sometimes only after spending some unfortunate time on the market. [Among the medications taken off the market recently: nighttime heartburn drug Propulsid (removed because of fatal heart

rhythm abnormalities), diabetes drug Rezulin (removed after causing liver failure), and irritable-bowel-syndrome treatment Lotronex (removed for causing fatal constipation and colitis). All three were taken off the market in 2000].

While it cannot be claimed that the present invention solves all of these problems, it does offer a broader and more comprehensive view both for a protein target and for the action of a drug compound on the cellular machinery. This is because the present invention provides comprehensive information about protein production in the cell, showing precisely how, when, in what order and in what amounts does the cell respond to the compound following administration. The target itself can be seen in the context of other proteins that are co-synthesized with it, before it, or after it; connections with other proteins can be identified; Similarly, the compound can be seen in the context of other compounds that elicit a similar response, allowing SAR and QSAR analyses to be carried out.

All in all, apart from the hitherto unavailable information that the present invention provides, the technique holds the important promise of both widening and shortening the drug development roadway by early removal of compounds from the pipeline, by providing a information about a candidate much earlier in the process, and by allowing more compounds and targets to enter this roadway. Thus, the present invention has the potential to produce many more drugs in shorter time and with smaller expenditure.

Another important application of the present invention is as a tool for process optimization, process control and quality control of protein production, either in bio-reactors using bacteria or cell culture, or else in cell free translation systems. In these situations, the present invention can provide indispensable information about the amounts of the target protein being produced, as well as on the precise structure of the proteome backdrop to this manufacturing, ensuring that the desired protein is produced in precisely the required environment. This level of control, unavailable today, can create a revolution in the way proteins and protein drugs are produced and certified. This can lead to new protein production methods that are easier to control than current ones.

The technology that is most like the present invention in the context of drug development is the technology of cell-based assays (see above), which has been making its mark in recent years. Cell based assays offer whole cell monitoring, via one or more measurable cellular characteristics – such as calcium intake, dye response, or the optical monitoring of several proteins simultaneously. Though this is better than an assay of a single receptor *in-vitro*, the technique lacks the global view of the cell that the present invention offers, as well as the sensitivity, comprehensiveness, immediacy, and generic characteristic of the present invention.

A piece of technology that is similar in spirit to this disclosure is PCT application WO 01/16375 to Schneider and Rubens (S&R). In the S&R method, real time sequencing of DNA or RNA is proposed using FRET technology. There are several main differences between the two methods: the present invention deals with proteins where S&R deals with nucleic acids; the present invention identifies proteins with the aid of a database where S&R proposes de-novo sequencing; the present invention provides dynamic information where S&R technology describes sequencing alone; the present invention targets the ribosome where S&R technology targets the DNA or RNA polymerase.

SUMMARY OF THE DISCLOSURE

LIST OF ABBREVIATIONS

PCR: Polymerase chain reaction, a method for *in vitro* amplification of DNA.

GPCR: Cell surface receptors that are coupled to heterotrimeric G-proteins (GTP-binding proteins).

MALDI-TOF: mass spectrometer that ionizes the sample with the technique of matrix-assisted laser desorption ionization, and measures masses using a time-of-flight mass analyzer.

LC-MS: a mass spectrometer that is directly coupled to a liquid chromatography column, and where ionization commonly is achieved with the electorspray method.
LC-MS-MS: a mass spectrometer of the LC-MS type where ions are further fragmented and the mass spectrum of the fragments is measured. Often used to identify sequence of tryptic peptides.
ICAT: a method of sample tagging that allows relative quantitation of proteins in two samples using LC-MS-MS.
Y2H: A method for detecting protein-protein interaction in-vivo in *Saccharomyces cerevisiae.*
FRET: Fluorescence resonance energy transfer. A method by which molecular distances of the order of few nanometers can be determined using appropriate fluorphores.
HTS: High throughput screening, a method used in drug discovery by which a large library of chemical compounds is assayed for binding to a specific receptor.
ADME-TOX: A set of parameters relevant to drug candidates which should be measured prior to clinical trials (Absorption, Distribution, Metabolism, and Excretion) and TOX (Toxicity)
R&D: Research and development

The method disclosed here shows how the ribosome, the protein manufacturing plant of the cell, can be tricked into broadcasting its activity. The method relies on several recent technological advances: in biology, a new understanding of the ribosome's structure; in material science, novel fluorophores with new capabilities; in optics, setups that can detect single molecules and nanometer-scale inter-molecular distances. This allows us to construct a setup where the ribosome signals out which tRNA is currently being processed, which mRNA codon is being read, or which amino acid is currently being added to the nascent protein. This signal can be captured and analyzed in seconds to reveal the protein's identity. The procedure can be carried out for hundreds of

ribosomes simultaneously, and so provide, for the first time in biology, a tool for dynamic monitoring of protein synthesis.

With such a tool, the effects of potential drugs can be assayed to determine the proteins that are being up- or down-regulated by a compound, as well as describing the sequence of events; protein pathways can be identified by interpretation of the changing translation patterns, noting which proteins are just beginning to be translated and which are ceasing to be translated; rare, hydrophobic, and heavy proteins can be identified at the same efficiency as any other protein; cell-free translation systems can be monitored in real-time and protein production processes can be monitored and optimized. All this can be achieved in seconds rather than days, with a single type of engineered cell that is as near to a natural, non-engineered cell as possible. This is in strong contrast with the multitude of specialty, heavily engineered cell types that are required today. Indeed, the present invention stands to become a "universal assay" for drug discovery and development on the one hand, as well as an indispensable tool for basic and applied scientific biological and pharmaceutical research and production.

The disclosure provides a method and device for measuring and monitoring protein synthesis by the ribosome in real time, *in-vivo* as well as in *in-vitro* translation systems. In the present invention, the ribosome is engineered to carry a donor fluorophore in a special configuration, and tRNA and/or amino acids are either engineered to carry acceptor fluorophores or else their natural fluorescent properties are utilized. As the ribosome mechanism reads the mRNA information, incorporates tRNA molecules and synthesizes a polypeptide chain, a light source illuminates the ribosome, exciting the donor fluorophores and thereby the acceptor fluorophores whenever these are within sufficient proximity of the corresponding donors. The resulting signals are detected by an appropriate optical setup and collected by a computerized analysis system. The resulting information is used as a key for database searching and identification of the protein being synthesized at that moment. This data can be tabulated to correlate with chemical or environmental effects on the cells being studied and so decipher

the functionality of chemical compounds; to enable elucidation of cellular proteomic mechanisms; to control protein production systems; and to study the factors effecting protein synthesis.

In order to better understand the disclosure, current technologies that the invention utilizes will now be briefly explained.

The FRET technique

FRET – Fluorescent Resonance Energy Transfer – has been known for over 50 years (De Angelis, D.A., *Why FRET over genomics?* Physiol. Genomics 1999, 31; 1(2): 93-9; Selvin, P.R., *The renaissance of fluorescence resonance energy transfer*, Nat. Struct. Biol. 2000 Sep;7(9):730-4; Kenworthy, A.K., *Imaging Protein-Protein Interactions Using Fluorescence Resonance Energy Transfer Microscopy*, Methods. 2001 Jul;24(3):289-96). The technology allows measurement of distances in the nanometer scale up to about 10 nanometers. The idea relies on a quantum-mechanical principle where, under suitable conditions, energy can be transferred between molecules without photon exchange. In FRET, a donor fluorophore is excited by incident light while an acceptor fluorophore is nearby. The emission spectrum of the donor must overlap with the excitation spectrum of the acceptor. In such a configuration, some of the energy is transferred from donor to acceptor without generation of photons. Thus, donor emission decreases and acceptor emission increases as the distance between them diminishes. The energy transfer efficiency obeys the relationship $E \propto [1+(R/R_0)^6]^{-1}$, where R is the distance between donor and acceptor and R0 is a constant that depends on the configuration and characteristics of acceptor and donor.

FRET has been recently used to monitor DNA hybridization, identify SNPs in solution, and monitor protein-protein interactions. Numerous ingenious variants of the technique have been used successfully both *in-vitro* and *in-vivo*. One common application is the dynamic monitoring of inter-molecular distances.

Indeed, it is becoming common nowadays to observe dynamics of molecular interactions, in real time, in living cells.

Single molecule detection

An important ingredient in the present invention is the detection of single molecule FRET signals. This is aided by the recent advances in single molecule technology (Schwille, P. and Kettling, U., *Analyzing single protein molecules using optical methods*, 2001, Current Opinion Biotech., 12:382-386; Weiss, S., *Fluorescence Spectroscopy of Single Biomolecules*, 1999, Science 283, 1676-1683). Using such technologies, single molecules have been observed and measured – in solid state and in biochemical environments, both immobilized and in solution, *in-vitro* (Deniz, A.A., et al., *Single-molecule protein folding: Diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor* 2, 2000, PNAS 97 (10), 5179-5184) and *in-vivo* (Harms, G.S. et al., Single-Molecule Imaging of L-Type Ca2+ Channels in Live Cells, 2001, Biophysical Journal 81, 2639-2646). The technology relies on confocal microscopy, aided by advanced optics such as two-photon excitation and fluorescence correlation spectroscopy (Heinze, K.G., Koltermann, A., and Schwille, P., *Simultaneous two-photon excitation of distinct labels for dual-color fluorescence crosscorrelation analysis*, 2000, PNAS 97 (19), 10377-10382).

Particularly interesting is the merging of single molecule technique with FRET (single-pair FRET – spFRET, cf (Ha, T. et al., *Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism*, 1999, PNAS 96, 893-898)). With this technology, protein interactions as well as protein folding have been measured. Confocal microscopy allows the excitation and measurement of signals from a localized region with volume of less than 10^{-15} Liter. With properly labeled proteins there is a low chance of finding more than one tagged molecule in this volume. Appropriate optical setup allows monitoring the real-time inter-molecular or intra-molecular distance (as well as rotational position) in real-time, *in-vivo*, at an appropriate sampling rate (milliseconds).

The ribosome and the mechanism of translation

It is appropriate to briefly review here the structure of the ribosome and the mechanics of translation, as have been revealed by recent work (Alberts, B., Johnson, A., Lewis, J., Raff. M., Roberts, K., and Walter, P., Molecular Biology of the Cell, 4th ed, 2002, Garland Science, N.Y.; Ramakrishnan, V., *Ribosome Structure and the Mechanism of Translation*, 2002, Cell 108 557-572; Schlunzen, F. et al, *Structural basis for the interaction of antibiotics with the petidyl transferase center in eubacteria*, 2001, Nature 413 814-821; Sytnik, A. et al, *Peptidyl Transferase Center Activity Observed in Single Ribosomes*, 1999, J. Mol. Biol. 285, 49-54; Nyborg, J., and Liljas, A., *Protein biosynthesis: structural studies of the elongation cycle*, 1998, FEBS letters 430, 95-99).

The ribosome itself is composed of two subunits, termed 30S and 50S (there are some differences between bacterial and eucaryotic ribosomes– henceforth in this discussion the bacterial ribosome will be presumed). The large unit is composed of a pair of large RNA molecules (5S and 23S), the small subunit of a single RNA molecule (30S). Each unit has several dozen small proteins attached to it (Alberts, B., Johnson, A., Lewis, J., Raff. M., Roberts, K., and Walter, P., *Molecular Biology of the Cell*, 4th ed, 2002, Garland Science, N.Y.). The ribosome reads the code on mRNA molecules and synthesizes the encoded protein through the mediation of tRNA molecules. The process is carried out in three steps: initiation, elongation and termination. For our purposes elongation is the most important step.

The ribosome uses an adaptor molecule – transfer RNA, or tRNA. These molecules are a special type of RNA. At their one end, they have the anticodon part that binds to the RNA codon. At their extreme other end, they carry the amino acid corresponding to that codon. Figure 1A shows a cartoon of the ribosome in action. Three tRNAs are seen in the E (exit), P (Peptidyl), and A (Aminoacyl) sites, with two additional tRNAs in standby. The heavy line indicates the mRNA being translated, and the dotted line represents the polypeptide being synthesized, tied into the Peptidyl position. The main steps of elongation are as

follows. Step 1: Codon recognition. A tRNA carrying an amino acid binds to a vacant A-site, while the nascent polypeptide is attached to the P-site. Step 2: Peptide bond creation. A new peptide bond is created and the polypeptide chain is moved to the A-site. Step 3: Translocation. The ribosome translocates a distance of 3 nucleotides with respect to the mRNA, the two tRNA units and the polypeptide chain. Step 4: the cycle repeats itself until a stop codon is reached.

The attachment of amino acids to the tRNA molecules is performed by the aminoacyl-synthetase enzyme family. Discharging, or de-alcylation, is performed by the ribosome, serving as a ribozyme (RNA enzyme).

Basic description of the PSM method

The method of present invention will now be described. The ribosome of a live bacterium, cell, or an *in-vitro* translation mechanism is suitably tagged with one or more donor fluorophores. An acceptor fluorophore is placed on some of the tRNAs and/or on some of the amino acids. In one preferred embodiment the natural fluorescence of tRNAs and/or amino acids is utilized. In another embodiment a fluorophore is attached to some tRNAs and/or amino acids. It is clear to anyone skilled in this art that the inverse configuration can also be utilized, that is, to tag the ribosome with acceptor fluorophores and the tRNAs and/or amino acids with donor fluorophores. Henceforth, for the sake of clarity and without sacrificing generality, only the first of these similar configurations will be described.

An optical apparatus detects an engineered ribosome by directing electromagnetic radiation of the required wavelength and energy onto the ribosome, thereby exciting the donor fluorophores that have been attached to it. The acceptor fluorophores on the tRNAs and/or amino acids, whether engineered or natural, respond to this energy with the FRET effect whenever donor and acceptor are in sufficient proximity, that is, when the tRNA or an amino acid passes near a donor tag. The fluorescence radiation emitted from the acceptor fluorophores is detected by the optical apparatus and the event is recorded by a computerized analysis station. Since only some of tRNA/amino acid species are

tagged, and not all of them, the resulting signal contains a sequence of detection events interspersed with periods of no detection. In computer science terminology this is described as a bit stream with zeroes and ones. The stream contains some uncertainty as to the number of bits in each field, as well as other elements of uncertainty. The signal is then correlated with a database of signals computed from a relevant database of protein sequences. The method disclosed herein describes how to utilize this signal to identify the protein that is most likely to have produced said signal.

The method is able to take into account the precise nature of the uncertainty in order to provide a scoring function that is able to quantitate the measure of likelihood, for each protein in the database, that this protein produced the signal that was measured.

In one embodiment disclosed herein the acceptor fluorophores are the naturally fluorescent tRNAs. An example of this is tRNAPhe which contains the highly modified Y base (wybutosine) in position 37 (Langlois, R., Kim, SH and CRA Cantor, *A Comparison of the Fluorescence of the Y Base of Yeast tRNA Phe in Solution and in Crystals*, 1975, Biochemistry 14: 2554-2558; Huang, K.H. and Cantor, C.R., Studies of 30 S *Escherichia Coli Ribosome Reassembly Using Individual Proteins Labeled with an Environmentally Sensitive Fluorescent Probe,* 1975, J. Mol. Biology 97, 423-441). Also well known are the fluorescent characteristics of the amino acids tryptophane and tyrosine. In this particular embodiment, only the donor fluorophores need to be specially engineered onto the ribosome in the appropriate position.

In another embodiment, tRNAs are tagged, and the ribosomal donor tags are placed at an appropriate position near the A, P or E site of the ribosome. The tRNA molecule is amenable for tagging, as it has non-hybridized loops with nucleotides that can easily be hybridized. tRNA molecules can also be modified in one of several ways to attain the required fluorescent characteristic. The tRNA tag should have a high binding rate on the one hand, yet allow the tRNA to be processed in the ribosome without interference. The ribosomal tag (donor, in this

example) can be placed either on a suitable location on the rRNA, using hybridization, or on one of the ribosomal proteins, using genetic engineering or antibodies, or one of numerous other tagging techniques. This tag can be placed either near the A, the P, or the E site. At the A-site, a FRET signal will mark docking of tRNA at the site. In one preferable embodiment, the tag is placed near the E-site to ensure that the tRNA is detected following the proofreading stage. In this embodiment, the detected FRET signal correlates with the sequence of tRNAs encoded for by the mRNA being translated.

In another embodiment, the tRNAs are not tagged. Instead, both donor and acceptor tags are attached to the ribosome. One tag ("marker") is small and static, while the other one ("lever") mimics a tRNA molecule and operates as a lever on the exiting mRNA. The lever is constructed to bind weakly to a specific nucleotide pair – say GC. Whenever the mRNA sequence contains a GC in the appropriate position, the lever goes "down". For other pairs, the lever stays "up". The FRET signal thus marks the advance of the mRNA with a bimodal response. This embodiment has the important benefits of providing a signal for every codon (rather than for blue codons only), and of not suffering from problems of proofreading. A diagram of this embodiment is shown in Figure 2. In this embodiment, the detected FRET signal correlates with the sequence of the mRNA being translated.

In another embodiment, the donor tag is placed near the peptide exit channel, and so is able to excite the acceptor fluorophores on the amino acids, either natural or engineered. In this embodiment, the detected FRET signal correlates with the amino acid sequence of the protein being synthesized.

In all these embodiments, several different tags can be used: Green Fluorescent Protein (GFP) and its derivatives, tetramethylrhodamine (TMR), Cypro dyes, and lanthanides. In one preferable embodiment quantum dots are used (Chan WC, Maxwell DJ, Gao X, Bailey RE, Han M, Nie S., *Luminescent quantum dots for multiplexed biological detection and imaging,* Curr Opin Biotechnol 2002 13(1):40-6, *Semiconductor nanocrystals as fluorescent biological labels,* Science. 1998 25;281(5385):2013-6).

These semiconductor nanoparticles can be engineered to provide fluorescent probes. Compared with conventional fluorophores, the nanocrystals have a narrow, tunable, symmetric emission spectrum and are photochemically stable. They can be engineered to be orders of magnitude brighter and more stable than the traditional fluorophores. This means that monitoring can be performed for minutes of even hours, in contrast with the traditional probes, that may bleach out after a few seconds. In one preferred embodiment, FRET architecture makes use of a donor-quencher pair rather than a donor-acceptor architecture. This is a more robust architecture, allowing the use of metal and other particles instead of fluorescent biomolecules. In this embodiment, the signal measured is in fact the donor signal, interrupted by periods of quenching.

In yet another preferred embodiment, a method is disclosed by which hundreds of ribosomes can be monitored simultaneously. This is important in order to ensure detection of rare proteins, such as proteins that have a very low copy number, sometimes less than one copy number per cell. The size of the ribosome (along the mRNA) is approximately 20 nM. Along the mRNA there are usually numerous ribosomes at various stages of translation. The distance between these ribosomes is about 30 nM. A typical mRNA may be attached to dozens of ribosomes, depending on the mRNA length. This means that it is sufficient to tag only about 10% of ribosomes. This can be achieved by limiting the tag concentration with respect to the number of ribosomes in the assay. A microscopy setup that has a resolution of about 100nM will allow monitoring of single ribosomes (cf. Sytnik, A., et al., *Peptidyl Transferase Center Activity Observed in Single Ribosomes*, J. Mol. Biol. (1999) 285, 49-54).

There are anywhere from 1000 to 20,000 ribosomes in a single cell, so that in order to identify single-copy number proteins 10% of that number should be imaged, i.e. 100-2000 ribosomes. This can be done either in one cell or on a cell culture with tens or hundreds of cells. An imaging device with pixel size of $100nM^2$ and 1000X1000 pixels will have a field of view of 100 square microns, which holds about 100 eucaryotic cells. There is a tradeoff between the number

of ribosomes monitored, the copy-number sensitivity, and the temporal response of the system. If only 1% of ribosomes are monitored, than a protein will be detected only once in 10 syntheses, which will take 10 times longer to detect than when 10% are monitored.

In one preferred embodiment, the optical setup includes an appropriate CCD camera with sufficient sensitivity and resolution. Other preferred embodiments use a photomultiplier tube-based imaging device. In yet another preferred embodiment, confocal microscopy is used to both excite the donor fluorophores and to detect the emitted acceptor FRET signals. The confocal setup will also assist in decreasing the amount of photobleaching, thereby increasing the total measurement interval. It will also assist in resolving the monitored ribosomes and separating the signals from the different ribosomes.

In another preferred embodiment, the detected PSM signals are used to identify the protein being synthesized with a computerized system that includes a protein database, identification software, and a scoring function. In this embodiment, the detected PSM signals are recorded by the computer and arranged into sequences of signals, one for each resolved ribosome. Each signal can be transformed into a sequence of zeroes and ones, amenable to computer analysis. In this embodiment, a protein database has been compiled to provide computed sequences that model the sequences observable with the disclosed method and device. A scoring function is provided together with a software module that is able to quickly compare the modeled sequences with the measured ones, and identify the database protein most likely to have produced the measured signals. This scoring function is constructed by carefully modeling the experimental setup, and takes into account the inherent measurement noise and uncertainty, for example in the proofreading step that may introduce changes in the duration of the synthesis step.

In another embodiment, the disclosed method is used as a system for high-throughput screening of chemical compounds in a pharmaceutical

application that tests thousands of chemicals for activity, toxicity and other parameters. In this embodiment, bacterial cultures or cell lines are prepared for PSM measurements by having their ribosomes and/or tRNAs, mRNAs or amino acids properly tagged. Then, in a system for high throughput screening appropriately adapted for protein synthesis measurements by the present invention, each sample is monitored while the chemical compound is administered. The results are recorded and analyzed by the computerized analysis system provided, and for each compound an analysis of the changes to protein synthesis that were caused by the administration of this compound are tabled. Screening software analyses these tables to detect which of the compounds has the desired effect according to prescribed conditions. Furthermore, once the compounds are screened and a relatively small number of them is considered suitable for continued development, the data collected by screening with the present invention is available for elucidating the pathways triggered by this compound and for understanding its function.

In another preferred embodiment, protein synthesis monitoring is used to identify protein pathways and protein-protein interaction maps. In this method, different chemical and environmental conditions are applied to a bacterial culture or to a cell line, and protein synthesis is monitored for each one. The protein synthesis patterns are indicative of the cellular pathways, and the protein synthesis data for the entire experiment is used to map the cells protein interaction, thereby enabling understanding of the intricate connections and functions of proteins in the cells.

In another preferred embodiment, the present invention is used to monitor proteins being produced by bacteria, yeast, cell lines, cell-free translation system protein production systems or any other protein production processes. Numerous biopharmaceuticals are thus produced, including drugs such as insulin, medical agents such as Erythropoietin, and other diagnostics and medical aids. The system disclosed here can be integrated into the production system, enabling

production yields and rates could be measured, controlled, and optimized in real time. This can lead to new protein production methods that are easier to control than the customary methods. This new monitoring capability can make a large impact on the quality, quantity and cost of protein production, as well as allowing tighter regulatory control of protein production.

To summarize, some of the advantages of the technology disclosed here are:

1. Providing protein synthesis dynamics information – identifying which protein is synthesized when, for the first time in the history of biological research.
2. Unprecedented sensitivity, down to very rare proteins – in a sense the present invention is like PCR for proteins, since even minute amounts can be detected.
3. Protein quantitation is obtained in essentially absolute numbers, at least as far as synthesis is concerned: how many proteins of a given type have been produced in the cell during a given time–frame.
4. The disclosed monitoring method can follow the chain of cellular events as they unfold via protein synthesis.
5. The disclosed method can complement other methods, such as protein tagging for monitoring protein localization and degradation.
6. The disclosed method can trigger specific protein studies once a protein is implicated as "interesting", to follow its life course.
7. The disclosed method can help elucidate protein pathways and interactions, and support analysis of protein function.
8. The disclosed method is generic – no need for special cell engineering for each application – one cell setup will be used always.
9. The disclosed method requires very mild cell engineering – all proteins are natural and the cell is as close as possible to a regular cell. This is as near as possible to "fly on the wall" monitoring.

10. The disclosed method can monitor protein production and assist in process optimization and control.
11. No need for presumptions or special preparations for a given application.

Technological comparison table

	The disclosed method	Proteomics technologies: 2D-Gels / MS	Protein Chips	HTS standard assays	Cell-based assays
Protein dynamics	Yes	No	No	No	Only for predefined proteins
Real time monitoring	Yes	No	No	No	Only cumulative
Protein sensitivity	Down to 1 copy number	>1000	>10,000	>10,000	Cumulative
Quantitation	Extremely accurate	Extremely inaccurate	No	No	Only cumulative
Analysis duration	Instantaneous	Days	1 day	Assay response time	Cell response time
Need for presumptions	No	No	Yes	Very much so	Very much so
Can follow chain of events	Yes	No	No	No	No
Generic	Yes	Yes	No	No	No
Mild intervention in cell	Yes	No (requires lysis)	No (requires lysis)	No (receptors are studied *in-vitro*)	No – proteins or cells are modified considerably

BRIEF DESCRIPTION OF THE FIGURES

FIGURE 1A describes the basic structure of a ribosome and the elongation cycle.

FIGURE 1B describes the tRNA molecule.

FIGURE 2 describes the embodiment of the ribosome tag strategy.

FIGURE 3 describes the embodiment of the tRNA tag strategy.

FIGURE 4 describes the elongation cycle of the translation mechanism.

FIGURE 5 describes the embodiment of the amino acid tag strategy.

FIGURE 6 describes the properties of a FRET pair and the dependence of the FRET effect on pair distance.

FIGURE 7 describes the optical setup for data acquisition.

FIGURE 8 describes in detail the optical setup at the sample scale.

FIGURE 9 describes optical and computational setup.

FIGURE 10 describes the analysis workflow in the system.

FIGURE 11 describes an embodiment of chemical library screening

DETAILED DESCRIPTION OF SEVERAL EMBODIMENTS

We now disclose in detail the new method for monitoring of protein synthesis. The method will be described in several variants and several applications by way of an example, and it should be recognized that the illustrated embodiments should not be taken as a limitation on the scope of the disclosure. The following sections will describe various aspects and details of the new invention.

Fluorescent or luminescent labeling of proteins and RNA

A large variety of fluorescent markers and probes are available from vendors such as Molecular Probes (Eugene, OR U.S.A.), which also provide instructions for use. Probes that attach to rRNA or tRNA loops can be engineered, as well as probes that attach to the ribosomal proteins. One widely used method to attach a fluorescent probe to a protein proceeds by reengineering the cell to produce the protein fused to a fluorescent protein, such as GFP (green fluorescent protein). Other donor fluorophores can be used by directly or indirectly attaching them to the preferred site on the ribosomal RNA or one of the ribosomal proteins.

Green fluorescent protein (GFP) includes a chromophore built of amino acids located in the center of the molecule. GFP excels in being photostable as well as having numerous variants with a choice of excitation and emission wavelengths (U.S. Patents 5,626,058 and 5,777,079; Herzenberg et al., Clin Chem. 2002 Oct;48(10):1819-27; Hailey et al., Methods Enzymol. 2002;351:34-49.) GFP can be attached to a ribosomal protein through the method of generation of a fusion protein, by well-known recombinant techniques as explained, for example; in Molecular Cloning, A Laboratory Manual, cold Spring Harbor Laboratory, Cold Spring Harbor, New York, chapter 17, 1989, herein incorporated by reference. A cell that is engineered to produce this fusion protein

will give rise to ribosomes that include the engineered ribosomal protein as required.

Numerous other methods exist for fluorescent labeling or dyeing a protein for fluorescent applications, as explained, for example in Allan, V.J. (ed), Protein Localization by Fluorescence Microscopy, A Practical Approach, Oxford University Press, herein incorporated by reference. In one embodiment, molecular biology techniques such as site-directed mutagenesis and unnatural amino acid mutagenesis (Anthony-Cahill et al., 1989, Trends Biochem Sci 14, 400) can be used to introduce cysteine and ketone handles for specific and orthogonal dye labeling of proteins (Cornish et al., PNAS 1994). Fluorescent-labeled nucleotide analogs can be used to site-specifically label RNA molecules. The large repertoire of molecular biology techniques and the ability to label many different sites on the macromolecule's surface offer great flexibility in dye labeling and thus in the generality and applicability of single molecule experiments (Weiss 1999, Science 283).

Preparation of Fluorescence-donor Ribosomes

Ribosome preparation requires attachment of a donor fluorescent or luminescent probe. The location of attachment varies with the strategy used (see below). Figure 1A shows a cartoon of bacterial ribosome structure with the larger (50S) subunit 10, smaller (30S) subunit 20, aminoacyl (A) site 50 where tRNAs dock initially, peptidyl (P) site 51 where the growing polypeptide chain is docked, and exit (E) site 52 where the tRNA is removed. Also shown are tRNAs that are undocked yet 40 and 41, the mRNA being decoded 30 and the nascent polypeptide chain being synthesized 45. The ribosome itself is made up of large folded rRNA chains with ribosomal proteins. The larger subunit contains two folded rRNAs, usually known as 23S and 5S. The smaller subunit contains one folded rRNA, 30S. On these rRNA chains more than 50 ribosomal proteins are docked. This complex structure enables a large variety of strategies for attaching fluorescent or luminescent probes. Depending on the strategy used, these

probes should be attached near the A site (50 in Figure 1A) or the E-site (52 in Figure 1A) when tRNA tagging is used; when rRNA tagging is used they should be located near the entry or exit channel of the mRNA; and when amino-acid tagging is used, the probe should be attached near the peptide exit channel.

Preparation or utilization of fluorescent acceptor tRNAs, mRNAs or amino acids

Figure 1B shows a tRNA molecule 60, with the anticodon loop 65, the amino acid arm 70, and a loaded amino acid 80. When tRNA is tagged, one needs to make sure that the tRNA continues to function normally in the processes of charging with amino acid, attaching to the elongation factors, and traveling through the ribosome. One appropriate solution is the use of Phe-$tRNA^{Phe}$-Fl^{8}, which forms a ternary complex and functions in protein synthesis (Johnson et al., 1982, 1986). The energy transfer acceptor species can be created by modification of GRP by covalent attachment of a single rhodamine dye to a ribose hydroxyl to form GTP-Rh (Watson et al., Biochmeistry 1995). In one preferable embodiment, use is made of the natural fluorescent properties of tRNA. It is well known that $tRNA^{Phe}$ is highly fluorescent, as it contains the highly modified Y base (wybutosine) in position 37 (Langlois, R., Kim, SH and CRA Cantor, *A Comparison of the Fluorescence of the Y Base of Yeast tRNA Phe in Solution and in Crystals*, 1975, Biochemistry 14: 2554-2558; Huang, K.H. and Cantor, C.R., *Studies of 30 S Escherichia Coli Ribosome Reassembly Using Individual Proteins Labeled with an Environmentally Sensitive Fluorescent Probe*, 1975, J. Mol. Biology 97, 423-441).

A similar phenomenon holds for amino acids, where it is known that Tryptophan and tyrosine are highly fluorescent and can be used without modifications. Phenylalanine is slightly less so. The fluorescence properties of these amino acids are summarized in the table below. The use of naturally fluorescent biomolecules is preferable as it requires less intervention and produces an engineered cell that is as similar to a wild cell as possible.

	Lifetime (Nanoseconds)	Excitation		Emission	
		Wavelength	Molar absorptivity	Wavelength	Quantum yield
Tryptophan	2.6	280	5,600	348	0.20
Tyrosine	3.6	274	1,400	303	0.14
Phenylalanine	6.4	257	200	282	0.04

Synthesis monitoring – ribosome tag strategy

Figure 2 shows another preferred embodiment, that of ribosomal tagging. In this method both donor and acceptor are attached to the ribosome. One tag ("marker") is small and static, while the other one ("lever") mimics a tRNA molecule and operates as a lever on the exiting mRNA. The lever is a modified tRNA molecule, with a "codon loop" and a relatively long and rigid arm. The lever is constructed to bind weakly to a specific nucleotide pair – say GC. In Figure 2A the mechanism is shown in "non-recognition" state. 100 and 102 are the ribosome subunits, 110, 111 and 112 are the A, P and E sites respectively. Only A and P are occupied in this diagram. tRNA-like lever 120, connected to large ribosomal subunit 102, does not match well the nucleotides on mRNA 104, passing between large subunit 102 and small subunit 100, in that the corresponding bases on the RNA chains are not complementary pairs (in the Watson-Crick sense). Therefore the lever stays relatively far away from marker 122 and a small or null FRET signal is obtained. In Figure 2B the mechanism is shown in a "recognition" state, where the tRNA-like lever 150 does match the nucleotides on mRNA 134. Therefore the lever moves relatively near to the marker 152 and a large FRET signal is obtained. 140-142 point to the A, P and E sites. If the lever is designed to recognize a "GC" pair on the mRNA, then whenever the mRNA sequence contains a GC in the appropriate reading frame, the lever goes "down". For other pairs, the lever stays "up". The FRET signal thus marks the advance of the mRNA with a bimodal response. This preferred

embodiment has the important benefits of providing a signal for every codon (rather than for only some of the codons).

Synthesis monitoring – tRNA tag strategy

tRNA-tagging involves placing a donor, say, on the ribosome and an acceptor on some of the tRNAs. The tRNA molecule is relatively amenable for tagging, as it has non-hybridized loops with nucleotides that can easily be hybridized. The tRNA tag should have a high binding rate (preferably above 90%) on the one hand, yet allow the tRNA to be processed in the ribosome without interference. The ribosomal tag (donor, in this example) is placed either on the rRNA, using hybridization, or on one of the ribosomal proteins, using genetic engineering or antibodies. This tag is placed either near the A, the P, or the E site. At the A-site, a FRET signal will mark docking of tRNA at the site. In one preferable embodiment the P or E sites are tagged. This ensures that the tRNA is identified after the stage of proofreading. At all sites, the geometry of the ribosome forms the basis for analyzing tag structure and docking. The tags should preferably be placed on the entry or the exit of the channel, to reduce interference with the synthesis process.

Figure 3 shows the basics of the elongation cycle. Step 1 – Codon recognition- is shown in Figure 3A. A tRNA molecule 260 carrying an amino acid 262 binds to a vacant A-site, while the growing polypeptide chain 270 is attached to the P-site where tRNA 264 carrying the amino acid 266 is positioned. Step 2, Peptide bond formation, is shown in Figure 3B. A new peptide bond is created between amino acid 266 and amino acid 262, and the polypeptide chain 270 is moved to the A-site. Step 3, translocation, is shown in Figure 3C. The ribosome translocates 3 nucleotides with respect to the mRNA, the two tRNA units and the polypeptide chain. Step 4: the cycle repeats itself until a stop codon is reached.

Recall that as a tRNA docks onto the ribosome-mRNA complex, it is bound with elongation factor EF-TU and a GTP molecule. Thus tagging should

be compatible with this complex, and, indeed, make use of the structure of this complex. The tag Phe-tRNA^{Phe}-Fl^{8} has been shown to be compatible with all these requirements.

Figure 4 shows one embodiment of tRNA tagging strategy. Ribosome small subunit 160 and large subunit 164 are attached to an mRNA 162 being processed. tRNA 170 is in A site, tRNA 172 in P site, tRNA 174 in E site, tRNA 176 has just been ejected out of the E site and tRNA 178 is already free of the ribosomal-mRNA complex. Donor fluorophore 190 is located on the large subunit just outside the mRNA exit channel (in proximity to the E site). tRNA 176 in this Figure is a tRNA^{Phe} with a fluorescent Y base 192 near the anticodon loop and therefore in close proximity to the donor fluorphore 190. The donor fluorophore emission spectrum overlaps the excitation spectrum of acceptor fluorphore 192. Radiation energy 194 is made to impinge on donor fluorophore 190 and excite it, causing part of the energy to transfer to acceptor fluorphore 192, exciting it and causing emission of FRET signal 196 of a lower energy and therefore lower wavelength. When the tRNA 192 is not a tRNA^{Phe}, a FRET signal will not be detected. Thus, as the ribosome processes the mRNA, an on-off signal is detected that corresponds to the tRNA being used and therefore allowing identification of the protein being synthesized.

Synthesis monitoring – Amino acid tag strategy

In this method the natural fluorescence of the amino acids tryptophan and tyrosine can be utilized. The donor fluorophore should be placed near the peptide exit channel. In one preferred embodiment a single donor fluorophore with suitable excitation and emission spectra is positioned near the peptide exit channel and excited with radiation of the appropriate wavelength. In another preferred embodiment two donors can be used to optimize the fluorescent characteristics of these two amino acids, and these donors can be excited by two different radiation sources optimized to the requirements of said donors. In Figure 5 an example of this type of embodiment is depicted. Ribosome small subunit 200 and large subunit 202 are attached to mRNA 204 being processed. tRNA

206 is docked in the A site and carries amino acid 220. tRNA 208 is docked in P site and carries amino acid 222. tRNA 210 is docked in the E site and its amino acid has already been transferred from the tRNA to the nascent peptide. Donor fluorophore 226 is attached to the ribosome in a location near the exit channel of the nascent peptide. Donor fluorophore 226 is illuminated with electromagnetic radiation of a frequency compatible with its excitation frequency. Emission frequency of donor fluorophore overlaps the excitation frequency of acceptor fluorphore 228. If the amino acid 224 is a tyrosine with natural fluorophore 228, than when it passes sufficiently near to the donor fluorophore 226 a FRET signal is generated. Thus when a tyrosine is being pushed out of the peptide exit channel a FRET signal is generated. When other amino acids are in this position a FRET signal will not be generated. Thus, as the ribosome synthesizes the protein, an on-off signal is generated that corresponds to the protein being synthesized, allowing it to be identified.

Synthesis monitoring – mixed strategies

It is clear to anyone skilled in the art that the embodiments described above are not mutually exclusive and any combination thereof can be used. For example, the natural fluorescence of both amino acids together with that of the Y base of $tRNA^{Phe}$ can be used by placing appropriate donor fluorophores both at the exit channel of the mRNA strand as well as near the exit channel of the nascent peptide, and illuminating the ribosome with several wavelengths. Although this means a more complex apparatus, the added confidence of detection may be worth the effort.

Tag chemistry

Several different tags have customarily been used for FRET: Green Fluorescent Protein (GFP) and its derivatives, tetramethylrhodamine (TMR), Cypro dyes, and lanthanides. An important new technology is that of quantum dots (Chan WC, Maxwell DJ, Gao X, Bailey RE, Han M, Nie S., *Luminescent quantum dots for multiplexed biological detection and imaging*, Curr Opin

Biotechnol 2002 13(1):40-6, and Bruchez M Jr, Moronne M, Gin P, Weiss S, Alivisatos AP, *Semiconductor nanocrystals as fluorescent biological labels*, Science 1998 25;281(5385):2013-6). These semiconductor nanoparticles can be engineered to provide fluorescent probes. Compared with conventional fluorophores, the nanocrystals have a narrow, tunable, symmetric emission spectrum and are photochemically stable. They can be engineered to be orders of magnitude brighter and more stable than the traditional fluorophores. This means that monitoring can be performed for minutes of even hours, in contrast with the traditional probes, which bleach out after a few seconds.

There has been extensive study of tagging strategies for tRNAs, ribosomal proteins, extension factors such as EF-TU, as well as related entities such as the aminoalcyl-synthetases. Quite a number of these biochemical entities are naturally fluorescent, such as tRNAPhe which contains the highly modified Y base (wybutosine) in position 37(Langlois, R., Kim, SH and CRA Cantor. *A Comparison of the Fluorescence of the Y Base of Yeast tRNA Phe in Solution and in Crystals*, 1975, Biochemistry 14: 2554-2558.

[32] Huang, K.H. and Cantor, C.R., *Studies of 30 S Escherichia Coli Ribosome Reassembly Using Individual Proteins Labeled with an Environmentally Sensitive Fluorescent Probe*, 1975, J. Mol. Biology 97, 423-441), as well as the amino acids tryptophan and tyrosine. In some cases, modified tRNAs have been tested to remain functional in protein synthesis (Watson, B.S. et al, Macromolecular Arrangement in the aminacyl-tRNA-Elongation Factor Tu-GTP Ternary Complex, *A Fluorescence Energy Transfer Study*, 1995, 34, 7904-7912). This fact makes it easy to choose an appropriate combination of donor-acceptor architecture, or alternatively fluorescence quenching architecture.

Tagging of spatially resolvable ribosomes

The diameter of the ribosome is approximately 20 nM. Along the mRNA there are usually numerous ribosomes at various stages of translation. The distance between these ribosomes is about 30 nM. A typical mRNA may be attached to dozens of ribosomes, depending on the mRNA length. As shall be

shown below, this means that it may only be necessary to tag about 10% of ribosomes. This means that a microscopy setup with resolution of about 100nM will allow monitoring of single ribosomes (cf. Sytnik, A., et al., *Peptidyl Transferase Center Activity Observed in Single Ribosomes*, J. Mol. Biol. (1999) 285, 49-54). There are anywhere from 1000 to 20,000 ribosomes in a single cell, so that in order to identify single-copy number proteins it is required to image 10% of that number, i.e. 100-2000 ribosomes. This can be done either in one cell or on a cell culture with tens or hundreds of cells. An imaging device with pixel size of $100nM^2$ and 1000X1000 pixels will have a field of view of 100 square microns, which holds 100 eucaryotic cells. There is a tradeoff between the number of ribosomes monitored, the copy-number sensitivity, and the temporal response of the system. If only 1% of ribosomes are monitored, than a single-copy protein will be detected only once in 10 syntheses, which will take 10 times longer to detect than when 10% are monitored.

All in all, current confocal micro-spectroscopy is certainly capable of performing this task.

Optical setup

Confocal fluorescent microscopy is the method of choice for single molecule studies, as explained in detail for example in (Schmidt et al., 1995). Confocal microscopes are available, for example, from Zeiss, Weesp, Netherlands, and from Leica Microsystems, Ede, Netherlands. Usually an inverted microscope is used with 100X objective, illuminated by a laser with suitable excitation wavelength and filter set to allow distinction between the excitation and emitted light. A cooled CCD (Charge Coupled Device) camera system such as available from Princeton Instruments, Vianen, the Netherlands allows the detection of individual fluorophores. The imaging device should, in one embodiment, have illumination spot radius of 400nM, pixel step size of 10nMX10nM and field of 1000X1000 pixels, so that the field of view is 100 square microns, which typically holds 100 eucaryotic cells and can resolve hundreds of ribosomes spaced a couple of microns apart from each other. Other

measurement strategies can also be used, such as for example fluorescence lifetime imaging. In one preferred embodiment, a photomultiplier tube-based scanning device or an intensified charge-coupled device can also be used, along with a suitable laser system to excite the donor fluorophores.

Figure 6 shows a diagram of FRET analysis. In 6A the spectral graphs of donor excitation 250, donor emission 252, acceptor excitation 254 and acceptor emission 256 are shown. Note the overlap between the spectral responses of donor emission and acceptor excitation. In Figure 6B the efficiency of energy transfer is shown as a function of donor-acceptor normalized distance (R_0 is equivalent to 50 in this chart).

Cell-free translation system

Cell free translation systems are well known for over 20 years. Recently a system with purified recombinant factors that has a high protein synthesis yield has been published (Shimizu et al., 2001). Kits and detailed instructions can be obtained from vendors such as Promega (Madison, WI). These systems are used for several applications, such as ORF validation and functional analysis of gene products. The systems contain ribosome-rich media with the required tRNAs and amino acids, and little or no mRNA. When mRNA is introduced, the ribosomes begin translation and proteins are produced. Often the proteins are produced radiolabeled. This enables the researcher to verify that the required proteins were in fact produced. The system disclosed here is easier to assemble *in vitro* than *in vivo*, since labeling techniques are more readily available and easier to implement. For example, immunolabeling of ribosomal proteins or rRNA epitopes is possible *in-vitro* but difficult to achieve *in-vivo*.

Translation of *in vitro* transcribed mRNAs: *In vitro* translation can be performed using kits such as the nuclease-treated rabbit reticulocyte lysate available from Promega, (Madison, WI). Before in vitro translation, cellular mRNAs are heated at 67 °C for 10 min to unfold secondary structures that would

eventually effect the efficiency of mRNA translation. Reactions are then assembled as recommended by the supplier in the presence of 20 mCi of [35S]methionine (ICN Biochemicals). Protein synthesis occurs during incubation at 30 °C. Customarily, the resulting proteins are purified and analyzed by radiolabeling. The procedure requires centrifugation, rinsing and immunprecipitation followed by separation on SDS-polyacrylamide gels. Following electrophoresis, gels are exposed to film or Phosphor B1 screens, and the bands corresponding to the synthesized protein verified.

With the present invention, the cell-free translation system could produce one protein or many proteins, and their identification and production rates could be measured, controlled, and optimized in real time. This can lead to new protein production methods that are easier to control than the customary methods of bio-production in reactors with bacteria, yeast or CHO-cells, for example. Since the *in vitro* translation system is fully controllable, and since it also allows co- and post-translational modifications, this method is an attractive alternative to current technologies.

A further advantage of the *in vitro* protein monitoring system disclosed here is that *in vitro* labeling techniques are more readily available and easier to implement than *in vivo* systems. For example, immunolabeling of ribosomal proteins or rRNA epitopes is possible *in vitro* but difficult to achieve *in vivo*. Antibodies against ribosomal proteins can be prepared as in Nadano et al., Jpn J Cancer Res 2000 Aug; 91(8). Antibodies can also be prepared against particular epitopes of rRNA. An appropriate fluorophore attached to such an antibody can form the required fluorescent donor for cell-free protein monitoring as disclosed herein.

Data acquisition and processing

Figure 7 describes one preferred embodiment for the optical setup for data acquisition, where laser 300 with appropriate wavelength and energy impinges a light beam 302 through dichroic mirror 304 and into an inverted confocal microscope 306, focusing on a movable slide 308 on which sample wells 310 are

located. The laser energy excites the donor fluorophores on tagged ribosomes in the sample, which transfer energy without photon transfer to the acceptor fluorophores on tRNA, amino acids, or the acceptor tags which have been attached to the ribosomal proteins or RNA loops, according to the strategy adopted. This energy transfer occurs when an acceptor tag is in sufficient proximity to a donor tag. When this FRET excitation occurs, the fluorescence signal emitted from the acceptor passes back through the microscope 306, through the dichroic mirror 304, to be collected by CCD camera 312. It is clear to anyone skilled in the art of single molecule fluorescence imaging that this configuration is shown by way of an example and does not limit in any way the scope of the disclosure. There are numerous variants and changes to the optical setup, including but not limited to changing the laser illumination to another type of illumination, changing the CCD camera to a photo multiplier tube, changing the filter from dichroic to another type and in other ways selecting different optical paths and configurations.

Figure 8 shows a close-up of the system shown in Figure 7 at the sample level. Laser illumination passes through the objective lens 320, which typically has a high numerical aperture such as N=1.4. The illuminated volume has the shape of a pinched cylinder, depicted by hyperbolic section 330. The movable microscope slide 322 supports well 326 in which the sample is enclosed. The imaged area is in the image plane 328, bringing to a minimum the volume of illuminated sample, and thereby bringing to a minimum the number of ribosomes imaged at any one time. The FRET signal to be detected returns along the same optical path between the lines 324 and into the objective 320, as described above.

In Figure 9 an overview of one preferred embodiment is presented, where illumination module 350 illuminates sample 354 through microscope 352, the resulting signals detected by detection module 356 and the resulting image transferred to computerized analysis station 360 which analyzed the images, records the FRET signals with their precise timing and coordinates in the image and correlates the resulting signal sequences with protein data from a database

362, to obtain protein identification data that is presented on the computer screen as well as written into database 362 for further analysis.

The optical data from the camera is received by the computerized analysis station as a sequence of images, preferably at a rate of at least 60 frames per second to ensure that the synthesis cycle, operating at a rate as high as 20 amino acids per second, is properly sampled. The software that analyses this image sequence operates by stages that are shown in Figure 10. In this paragraph the term "module" is always taken to mean a software or hardware module, or combination thereof. The image sequence 400 is first received by recording module 402. An image 403 is processed by preprocessing module 404 that identifies putative signals in the single image. These signals are output as a list of coordinates 405 to signal sequencer module 406 that tracks the signals and clusters them into a list of separated signal sequences 407. In this process random or otherwise unmatchable signals are filtered out. A signal sequence is preferably of the form $S = (t_1, x_1, y_1, s_1), (t_2, x_2, y_2, s_2), \ldots, (t_n, x_n, y_n, s_n), \ldots$. The signal sequencer module 406 updates this list of sequences S and each updated sequence is sent to sequence interpolator 408 which transforms the sequence S into a bit-stream 409 of zeroes and ones, estimating the number of non-signals between each signal pair in the sequence as well as estimating the number of consecutive signals in each signal field, since a single signal may correspond to more than one ribosomal cycles. This transformation is performed by changing analog signals into digital data, according to a sampling algorithm as previously described. The bit stream 409 is sent to database matcher module 410 that retrieves a list 411 all protein sequences that putatively match signal 409. A scorer module 412 determines, for each candidate protein sequence, the probability that this protein is responsible for the observed signal sequence. The list 413 of scored protein candidates is analyzed and every protein 415 with a probability higher than a predetermined probability threshold is displayed on the screen. This protein is also recorded to the database 420.

Database interrogation and protein identification

In this section the identifiability of a protein from the data obtained by use of the present invention is analyzed. Let's consider first the case of amino acid tagging. Let us suppose that *k* of the 20 amino acids can be marked as "blue" and *20-k* as "red". Assume that a subsequence of total length *l* (both blue and red) has been read. Assume uniform distribution of amino acids. In this case, $\alpha = k/20$ of the *l* residues can be expected to be blue, so that there would be $p = \alpha l$ blue residues and *l-p* red residues. The number of possible sequences would then be $N = \binom{l}{p} = \binom{l}{\alpha l}$. This number is maximal when *k* = 10 ($\alpha = 1/2$). A realistic assumption is that for confident identification of a protein the random hit probability should be better than $1:10^6$. This is because there are at most several hundred thousand proteins for any given organism. Thus, for confident protein identification, the formula above connects the amino acid coloring scheme with the length *l* of a polypeptide chain required for confident identification. The following table provides example values:

Number of "blue" amino acids	Minimal number of residues (*l*) required for confident identification	Synthesis time in bacteria (seconds)
1	80	4.0
3	42	2.0
5	28	1.4
6	26	1.3
10	23	1.06

Thus it is beneficial to partition the amino acids as equally as possible. It is also clear that protein identification can be obtained in under 2 seconds in bacteria (where synthesis rate is approximately 20 AAs a second). Protein synthesis in eucaryotes is about an order of magnitude slower.

As for ribosome-tagging, if the tag marks about one-quarter of the codons, then this is equivalent to marking 5 AAs out of the 20, so that a 28-residue chain is required for confident protein identification.

Applications for cell-based assays

This example describes how the methods disclosed herein can be used for the screening of a large library of chemical compounds to determine their efficacy or their potential for use as drugs. High throughput screening is a method developed since the late 1980s. Today systems are available that can screen up to a million compounds in one day. High throughput screening requires an assay to be devised that is compatible with the screening instrument, an assay that enables quick rejection of most of the compounds as irrelevant, and approves only a small fraction for continued research. The present invention is suitable for a very thorough and informative assay, as explained above, in the sense that it provides information not only concerning a single protein target that has been suspected to be related to the disease for which a drug is sought, but provides, for every compound tested, information about the full spectrum of proteins that the compound induces the cell to produce. Thus, functional activity of a compound can be usefully studied by subjecting it to protein synthesis monitoring assay as disclosed herein only once for a given organism or cell-line. The strategy suitable for screening of a chemical compound library with the system and method disclosed herein is depicted in Figure 11. First, a cell line with tagged ribosomes is cultured and placed in a multiwell plate 450. This can have a 96 well plate format, a 384 well plate format or any other format suitable for standard automatization. The wells 452 in the plate need to be optically amenable for microscopy as shown in Figures 7 and 8. A robot 454 administers one compound out of the library being screened into each well and protein synthesis analysis is performed by a protein synthesis monitoring system 456. A suitable sampling regime should be adopted, for example a protein synthesis monitoring measurement for 30 seconds every 10 minutes for a total of one hour. The list of proteins synthesized by the cell during the sampling period for the particular well

is stored in screening database 458. A sampling regime as described allows screening of 20 compounds per hour per protein synthesis measurement unit. In this way, screening of a library with one million compounds can take one year with 5 protein synthesis monitoring units or one month with 50. Obviously many alternative regimes can also be used.

Though seemingly slow compared to other screening methods, the method disclosed here has numerous attractive advantages. First, the library needs to screened only once. Thereafter, focused screening targeted at a specific receptor or disease model can be done *in-silico*, with the computer alone, analyzing the protein synthesis data collected. Second, only a tiny amount of compound is required since the library is screened only once. Most importantly, the data collected is orders of magnitude more informative and relevant to further research than single receptor binding assays or even cell-based assays. Thus, choice of drug leads and further development of drugs from these leads becomes much more of a research program and less of a guesswork. In this way, over a period of several months, a hitherto unobtainable amount of critical information is compiled about each of the compounds in the library. This is in contrast with the information obtained after a full run of customary screening. For example, a customary receptor-binding assay produces only one bit of information for each compound - "binds" or "doesn't bind". This produces virtually no information, for almost the entire library. In fact, screening with the present invention is not so much a screening system (that rejects most of the candidates), as much as a functional validation system for entire chemical libraries.

Other applications for pharameceutical research and drug discovery, development and manufacturing

The present invention can help validate the type, state, specialization condition etc. of cells, and can even help classifying live cells or bacteria. When a cell is subjected to protein synthesis monitoring analysis, the spectrum of proteins that are produced is a very strong characteristic of the cell, showing whether the cell is in a state of toxic shock, is in a certain stage of the cell cycle,

is apoptotic, is cancerous, is infected by a virus etc. This opens a whole gamut of applications for protein synthesis monitoring in pharmaceutical development, in basic research, in medical research, and even clinical applications for diagnosis and prognosis.

Another important application of protein synthesis monitoring is as a tool for process optimization, process control and quality control of protein production, either in bio-reactors using bacteria or cell culture, or else in cell free translation systems. In these situations, the present invention can provide indispensable information about the amounts of the target protein being produced, as well as on the precise structure of the proteome backdrop to this manufacturing, ensuring that the desired protein is produced in precisely the required environment and in the right amounts. This level of control, unavailable today, can create a revolution in the way proteins and protein drugs are produced and certified. This can lead to new protein production methods that are easier to control than current ones.

Applications for basic biological research

This example describes how the methods disclosed herein can be used for basic biological research. One important field of research is the understanding of protein and metabolic pathways in the cell. Numerous methods have been suggested and several technologies have been developed to try and decipher the complex net of protein interactions and pathways of bacteria and cells. These include protein-protein interaction mapping from assays, from computerized analysis of research papers, and from clever setups as described in (Shen-Orr, S.S., Milo, R., Mangan, S., and Alon, U., Network motifs in the transcriptional regulation network of Escherichia coli. Nat Genet 2002, 31(1): 64-8). With the present invention, a large step forward in pathway elucidation is possible. Since the response of the cell to an external stimulus is instantly revealed by the monitoring of protein synthesis, the following methodology is suitable for discovering the pathways. As described in the screening application above and in Figure 11, cells are cultured and placed in a suitable screening setup. The cells are then subjected to various external stimuli, such as temperature changes,

chemical deprivation (such as phosphate starvation), administration of toxic agents, attack by phages or virii and any other stimulus that is thought to be relevant to the pathway being investigated. The cells or bacteria are studied using the present invention for an appropriate duration before or following the stimulus, and the response of the protein synthesis apparatus of the cell is tabled. A computerized analysis module analyzes the different responses, including the sequentiality of protein synthesis, and determines causality relationships between them. By virtue of having numerous different stimuli generating numerous different responses, pathway elements that are constant can be determined as explained, for example, in Rosen et al., FEMS Microbiol. Ecol. 2001, and Hecker et al., Int. J. Med. Microbiol 2000. The dynamic information of cellular response to external stimulus is an extremely important probe of cellular behavior, that has hitherto been impossible to measure.

Other applications

In view of the large number of possible applications and embodiments of the present disclosure it should be recognized that the illustrated embodiments are only particular examples and should not be taken as a limitation on the scope of the disclosure. Some of the possible additional applications which are clearly enabled by the present invention are clinical applications, diagnostic applications, production of food, cosmetics, and other bioproducts, military applications concerning biological warfare, and many more.

We claim:

1. A method for detecting protein synthesis in a protein synthesis system, the method comprising:

Providing a marker for protein synthesis in the system, said marker causing electromagnetic radiation to be emitted;

Detecting said electromagnetic radiation emitted from the system; and

Analyzing said emitted radiation to determine protein synthesis activity in said system.

2. A method according to claim 1 where the system is a bacterium or bacterial culture.
3. A method according to claim 1 where the system is a cell or a cell-line or a cell culture.
4. A method according to claim 1 where the system is a cell-free protein translation system (in-vitro translation system)
5. A method according to claim 1 where one or more of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in the system are artificially adapted to provide said marker.
6. A method according to claim 1 where said marker comprises at least a portion of one or more of unaltered ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in the system.
7. A method according to claim 1 where the measured signal is a FRET (Fluorescence resonance energy transfer) signal.
8. A method according to claim 7 where the measured signal is obtained by energy transfer from one of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in the system to

another of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in the system.

9. A method according to claim 1 where the signal is a fluorescent signal.
10. A method according to claim 1 where the system is irradiated with electromagnetic radiation.
11. A method according to claim 1 where the emitted signal is measured with a microscope.
12. A method according to claim 1 where the system is treated so that only one ribosome would be measured.
13. A method according to claim 1 where the system is treated so that many ribosomes would be measured simultaneously.
14. A method according to claim 1 where the analysis includes identification of the protein or proteins being synthesized at a given time.

15. An apparatus for measuring protein synthesis by a protein synthesis system, said apparatus comprising:

a container for containing a plurality of components for the system, wherein at least one component is capable of emitting electromagnetic radiation due to protein synthesis activities;

a detection system to measure emitted radiation from the system; and

a computational device to analyze said emitted radiation and determine the protein synthesis activity in said system.

16. An apparatus according to claim 15 where the system is a bacterium or bacterial culture.
17. An apparatus according to claim 15 where the system is a cell or a cell-line or a cell culture.
18. An apparatus according to claim 15 where the system is a cell-free protein translation system (in-vitro translation system)
19. An apparatus according to claim 15 where said component that is capable of emitting electromagnetic radiation includes one or more of ribosomes,

ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals.

20. An apparatus according to claim 15 where said component that is capable of emitting electromagnetic radiation includes one or more of adapted ribosomes, adapted ribosomal RNA, adapted ribosomal proteins, adapted tRNAs, adapted amino acids, or other adapted biochemicals in the system.
21. An apparatus according to claim 15 where the measured signal is a FRET (Fluorescence resonance energy transfer) signal.
22. An apparatus according to claim 21 where the measured signal is obtained by energy transfer from one of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in the system to another of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in the system.
23. An apparatus according to claim 15 where the signal is a fluorescent signal.
24. An apparatus according to claim 15 where the radiation source is a laser.
25. An apparatus according to claim 15 where the detection system includes a microscope.
26. An apparatus according to claim 15 where only one ribosome is measured.
27. An apparatus according to claim 15 where many ribosomes are measured simultaneously.
28. An apparatus according to claim 15 where the analysis includes identification of the protein or proteins being synthesized at a given time.

29. A method for analyzing a chemical compound library, said method comprising:

Administering each of the compounds to a protein translation system

Measuring said system's response using the apparatus of claim 15 and/or the method of claim 1

Analyzing the measurement's results to provide information about said compound.

30. A method according to claim 29 where said system is a bacterium or bacterial culture.

31. A method according to claim 29 where said system is a cell or a cell-line or a cell culture.

32. A method according to claim 29 where said system is a cell-free protein translation system (in-vitro translation system)

33. A method according to claim 29 where one or more of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in said system are artificially adapted to optimized their optical properties.

34. A method according to claim 29 where the measured signal is created by natural optical properties of one or more of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in said system.

35. A method according to claim 29 where the measured signal is a FRET (Fluorescence resonance energy transfer) signal.

36. A method according to claim 29 where the measured signal is obtained by energy transfer from one of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in the system to another of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in said system.

37. A method according to claim 29 where the signal is a fluorescent signal.

38. A method according to claim 29 where the system is irradiated with radiation of a laser.

39. A method according to claim 29 where the emitted signal is measured with a microscope.

40. A method according to claim 29 where the system is treated so that only one ribosome would be measured.
41. A method according to claim 29 where the system is treated so that many ribosomes would be measured simultaneously.
42. A method according to claim 29 where the analysis includes identification of the protein or proteins being synthesized at a given time.

43. An apparatus for analyzing a chemical compound library, said apparatus including a well array plate, a robot for placing a protein synthesis system into the wells, a robot for administering chemical compounds into said wells, an automatic plate handling system, and an apparatus according to claim 15 to analyze protein synthesis by said system.
44. An apparatus according to claim 43 where said system is a bacterium or bacterial culture.
45. An apparatus according to claim 43 where said system is a cell or a cell-line or a cell culture.
46. An apparatus according to claim 43 where said system is a cell-free protein translation system (in-vitro translation system)
47. An apparatus according to claim 43 where one or more of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in said system are artificially adapted to optimized their optical properties.
48. An apparatus according to claim 43 where the measured signal is created by natural optical properties of one or more of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in said system.
49. An apparatus according to claim 43 where the measured signal is a FRET (Fluorescence resonance energy transfer) signal.
50. An apparatus according to claim 43 where the measured signal is obtained by energy transfer from one of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in the

system to another of ribosomes, ribosomal RNA, ribosomal proteins, tRNAs, amino acids, or other biochemicals in said system.

51. An apparatus according to claim 43 where the signal is a fluorescent signal.
52. An apparatus according to claim 43 where the system is irradiated with electromagnetic radiation.
53. An apparatus according to claim 43 where the emitted signal is measured with a microscope.
54. An apparatus according to claim 43 where the system is treated so that only one ribosome would be measured.
55. An apparatus according to claim 43 where the system is treated so that many ribosomes would be measured simultaneously.
56. An apparatus according to claim 43 where the analysis includes identification of the protein or proteins being synthesized at a given time.

57. A method for determining cellular protein pathways, said method comprising:

Selecting a cellular or bacterial culture

Placing said culture in several sample containers

Subjecting said culture to a different condition in each of said containers

Measuring protein synthesis in each of said containers according to the method of claims 29-42

Analyzing protein expression patterns in all containers to determine protein pathways.

58. An apparatus for determining cellular protein pathways, said apparatus comprising:

Cellular or bacterial culture selection means;

Sample container means to transfer and contain said cultures;

Means for subjecting said culture to a different condition in each of said containers; and

Protein synthesis apparatus according to the apparatus in claims 43-56.

59. A method for ribosome labeling to allow protein synthesis monitoring, said method comprising:

Selecting a fluorescent probe;

Selecting a location on at least one of a ribosomal RNA or on a ribosomal protein according to at least one of a characteristic of said probe or a characteristic of at least one of said ribosomal RNA or said ribosomal protein; and

Attaching said probe to said location.

60. The method of claim 59, wherein said selecting said fluorescent probe is performed according to at least one of a suitable excitation or emission property of said probe.

61. A method for protein production monitoring, said method comprising:

Selecting a protein synthesis system for PSM analysis;

Selecting a fluorescent probe;

Selecting a location on at least one of a ribosomal RNA or on a ribosomal protein according to at least one of a characteristic of said probe or a characteristic of at least one of said ribosomal RNA or said ribosomal protein;

Attaching said probe to said location to perform PSM; and

Analyzing signals from said probe to monitor the protein synthesis system.

62. A method for detecting protein synthesis in a protein synthesis system, the method comprising:

Providing a marker for protein synthesis in the system, said marker having a label;

Attaching said marker to at least one component of the system; and

Detecting said label to determine protein synthesis activity in the system.

Abstract:

A method and a device is disclosed for monitoring the synthesis of proteins by the ribosome in real time, *in-vivo* as well as in *in-vitro*. The ribosome is engineered to carry a donor fluorophore, and tRNA and/or amino acids are either engineered to carry acceptor fluorophores or else their natural fluorescent properties are utilized as acceptors. As the ribosome mechanism processes the mRNA and tRNA molecules and synthesizes a polypeptide chain, a light source illuminates the ribosome, exciting the donor fluorophores and thereby the acceptor fluorophores whenever these are in sufficient proximity to a donor. The resulting signals are detected and used as a key for database searching and identification of the protein being synthesized. The resulting data can be tabulated to interpret activity of candidate drugs, elucidate cellular proteomic mechanisms, allow controlled production of proteins, and in general provide a novel, important piece of information central to the solution of the cellular puzzle.

FIGURES

FIGURE 1A

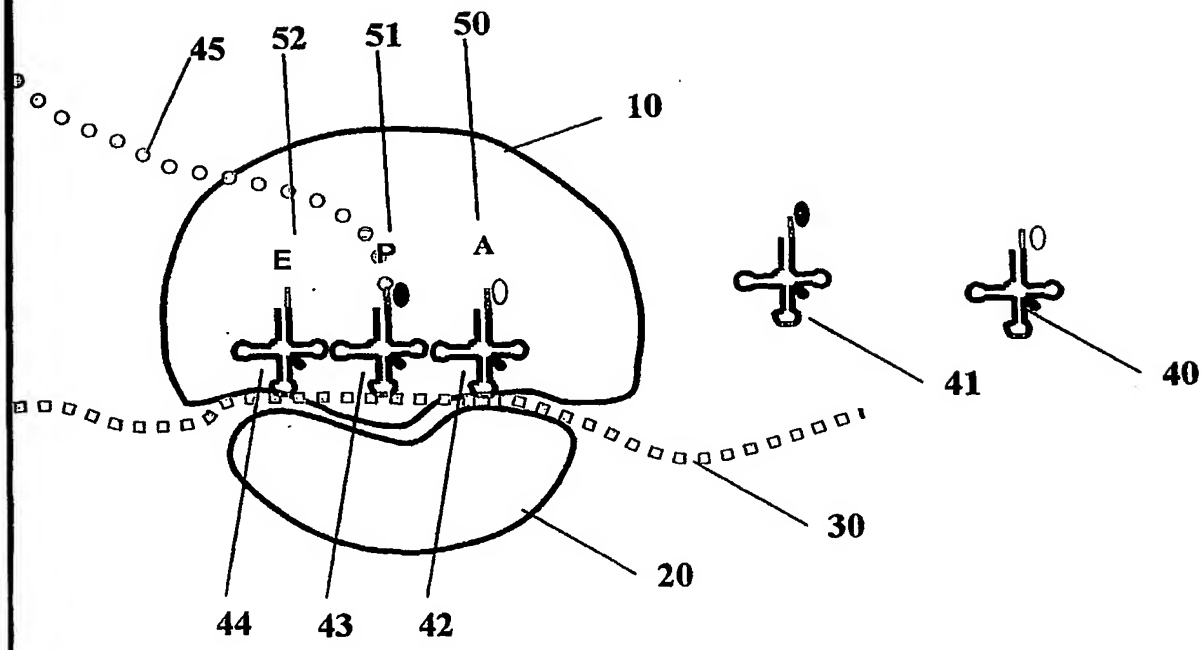


FIGURE 1B

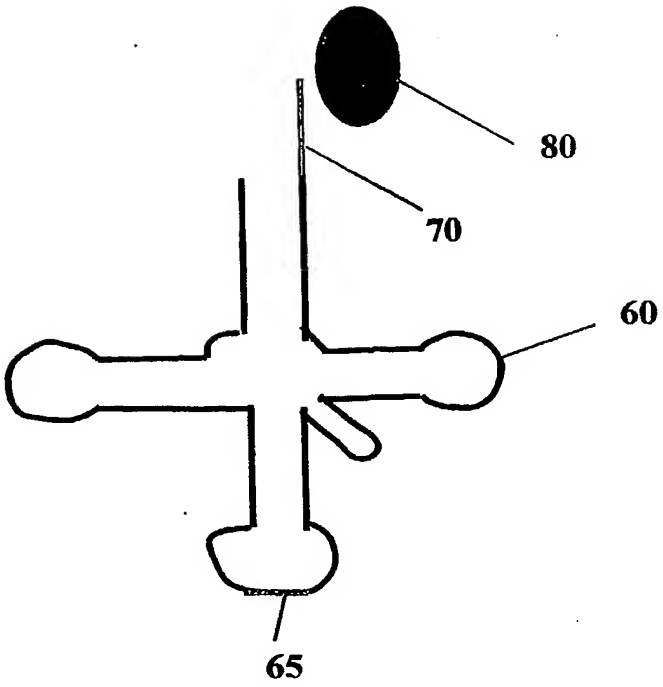


FIGURE 2

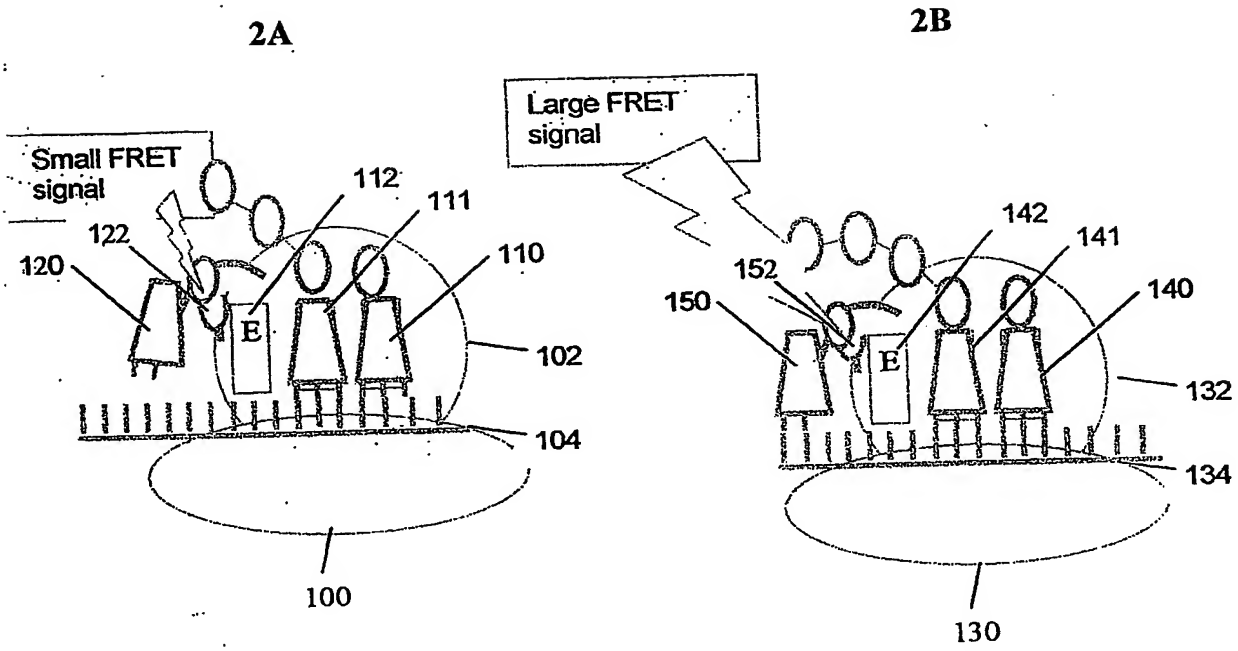
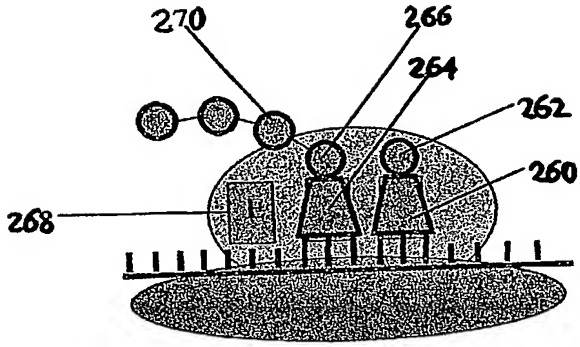
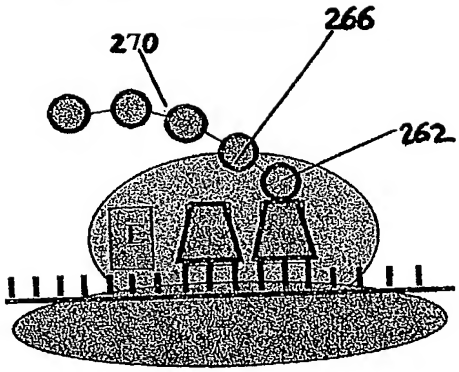

2A
2B
Small FRET signal
Large FRET signal
112
111
122
120
110
E
102
104
100
142
141
152
150
140
E
132
134
130

FIGURE 3

3A



3B



3C

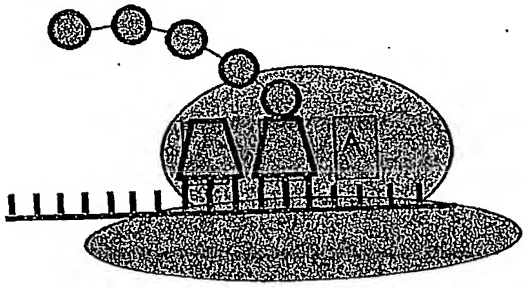


FIGURE 4

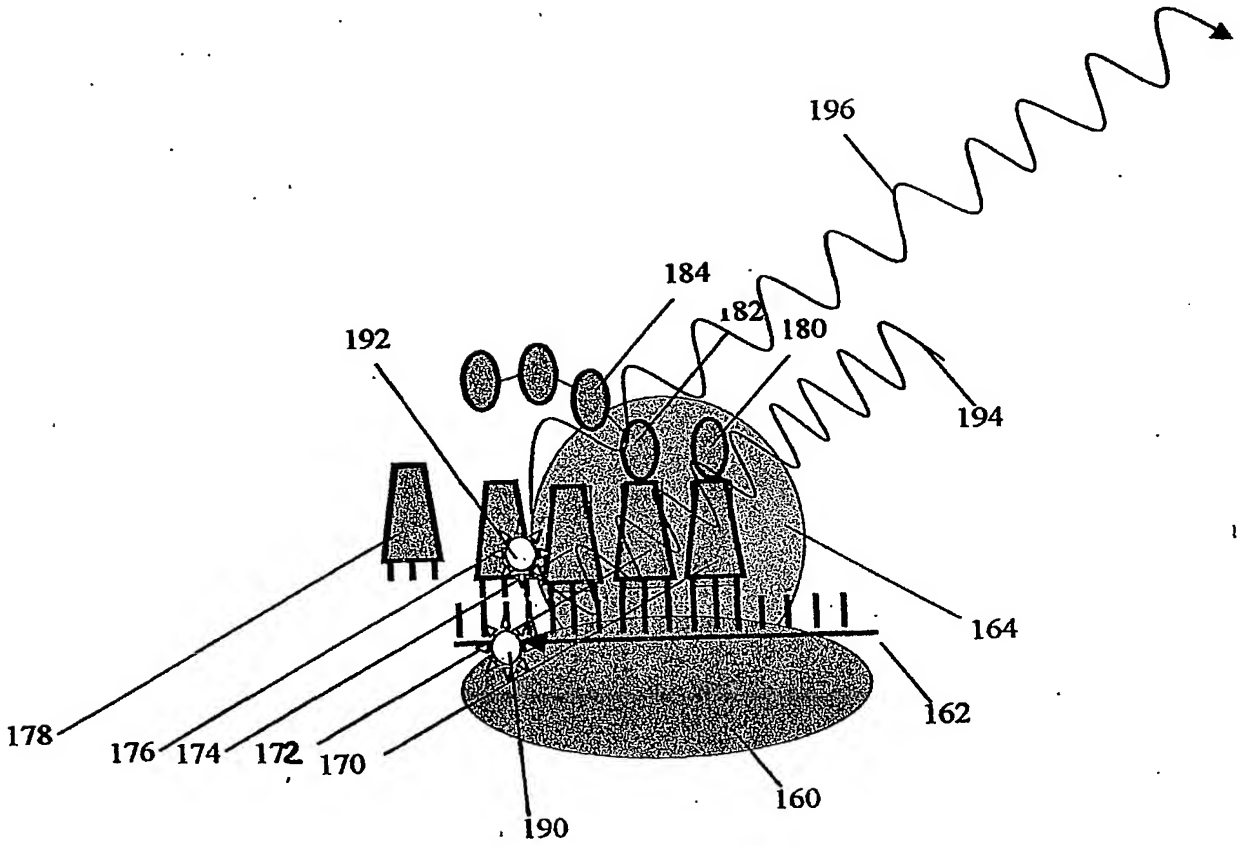


FIGURE 5

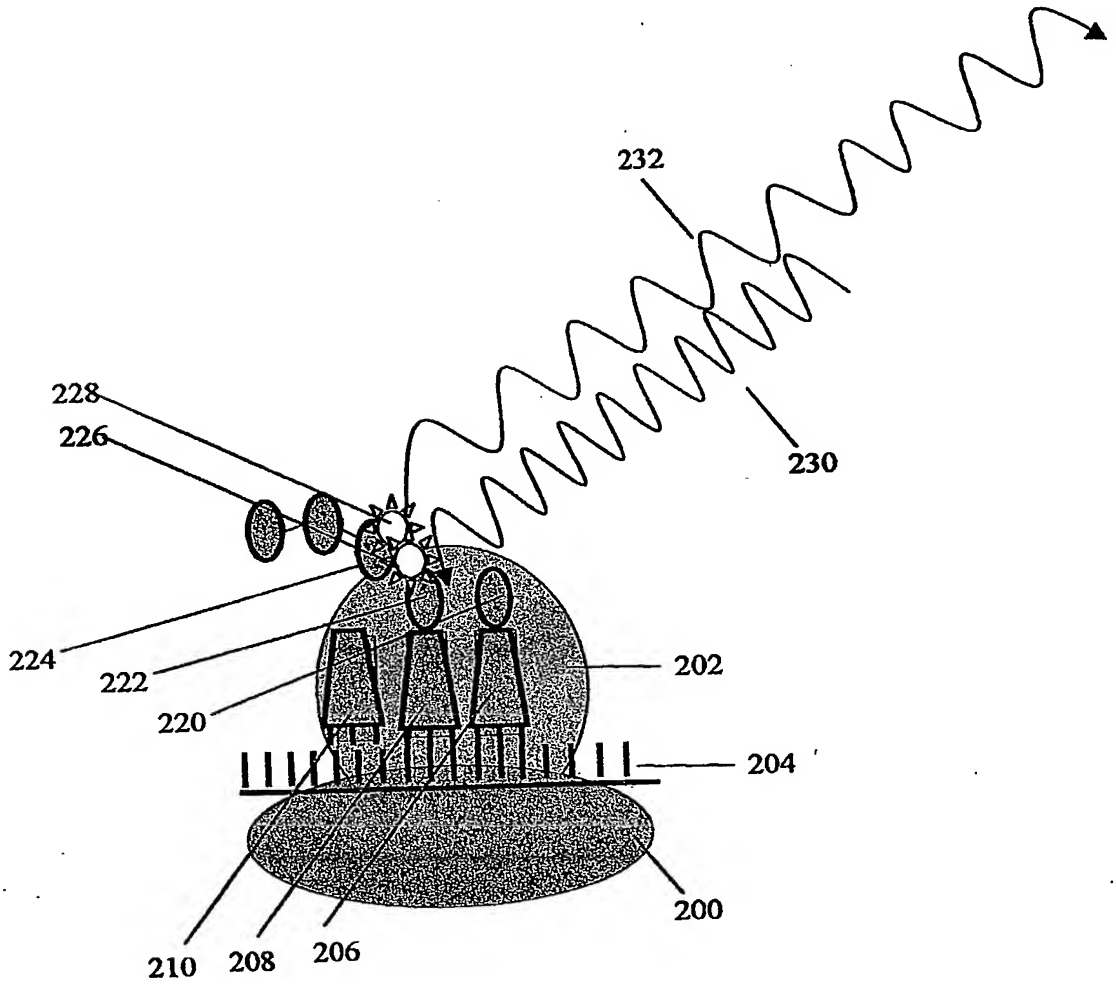
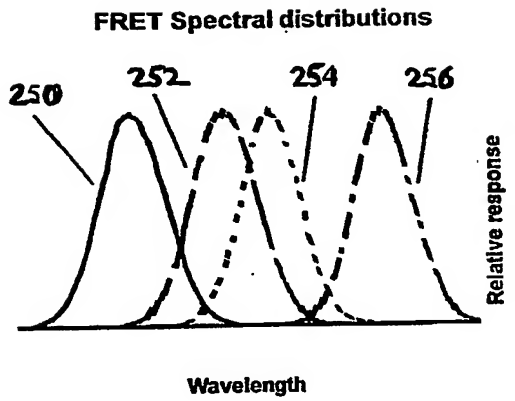
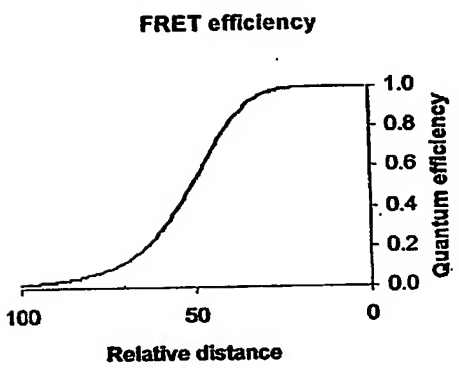


FIGURE 6



6A



6B

FIGURE 7

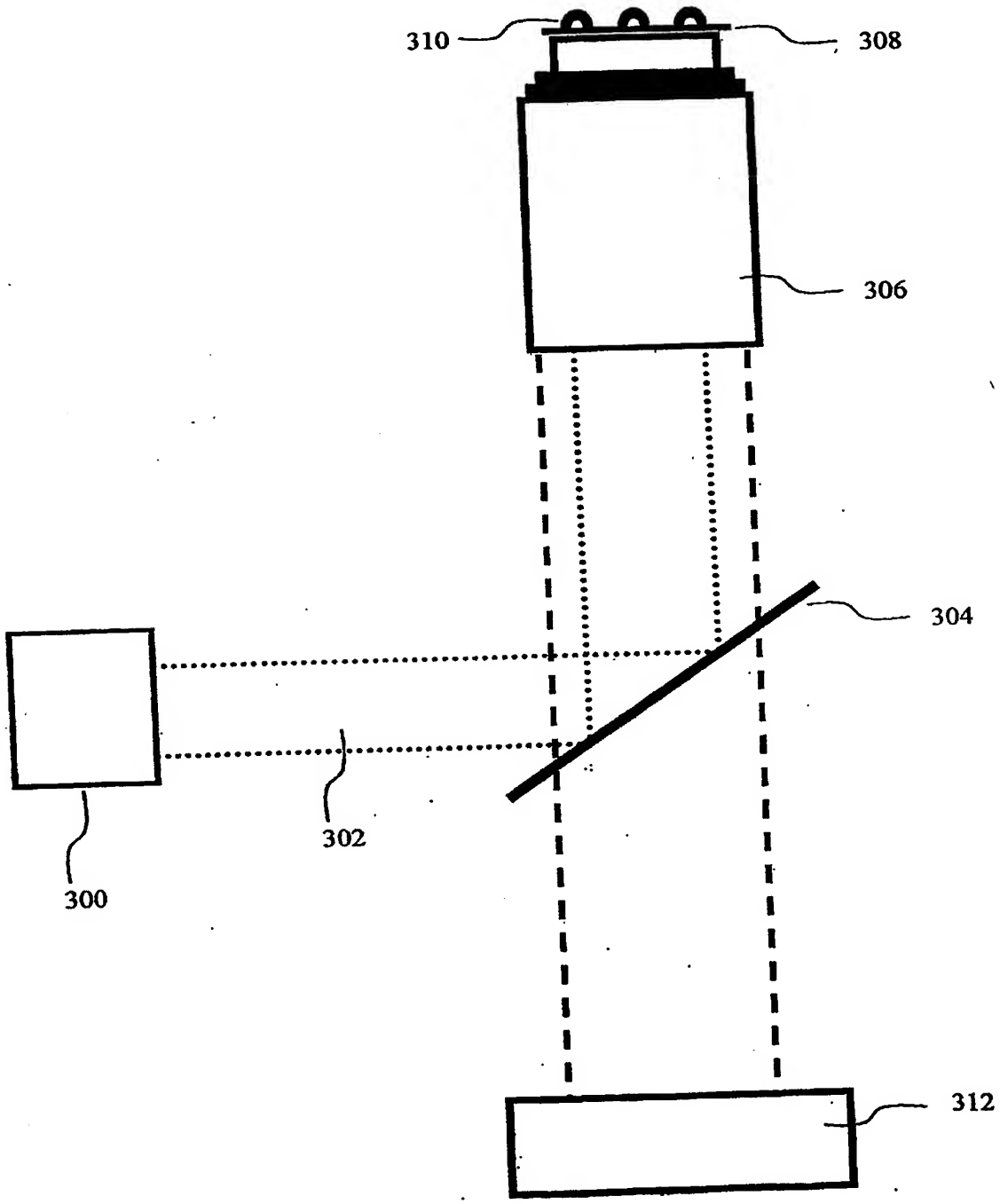


FIGURE 8

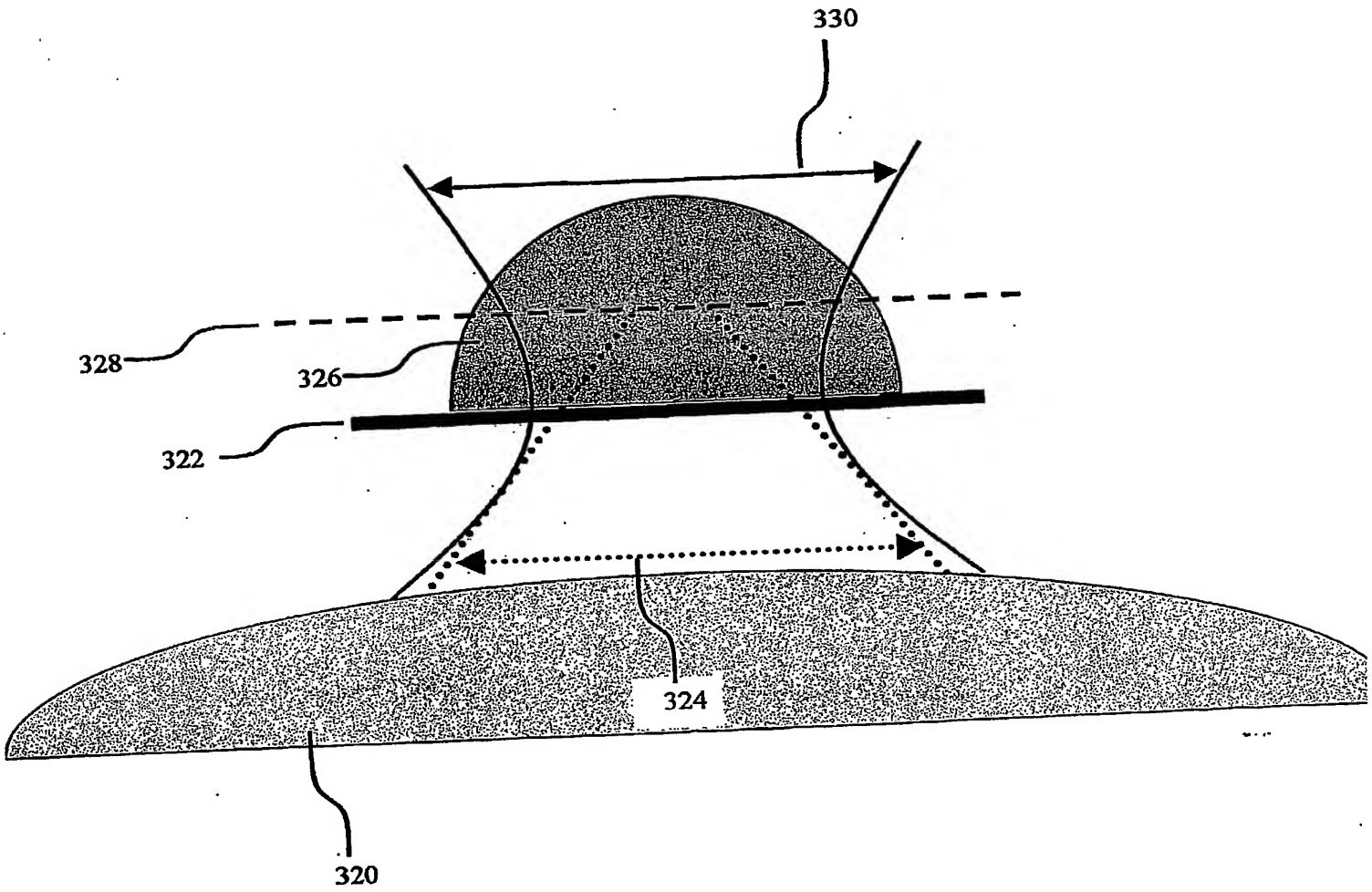


FIGURE 9

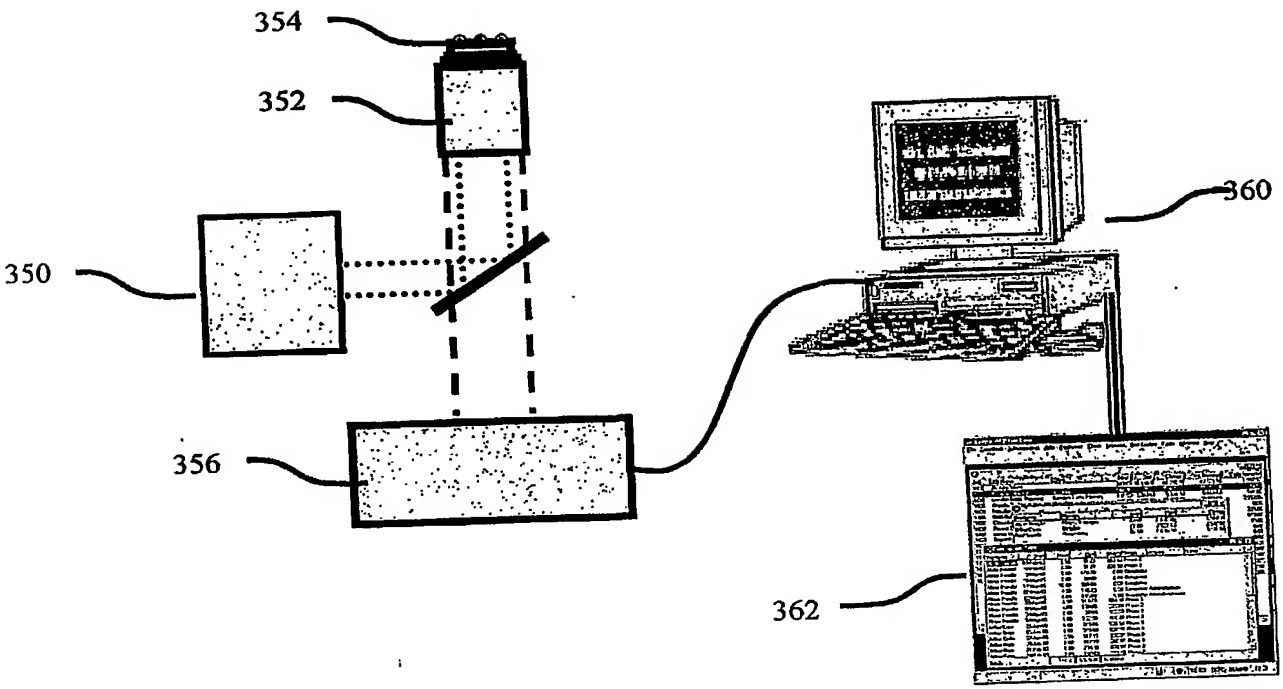


FIGURE 10

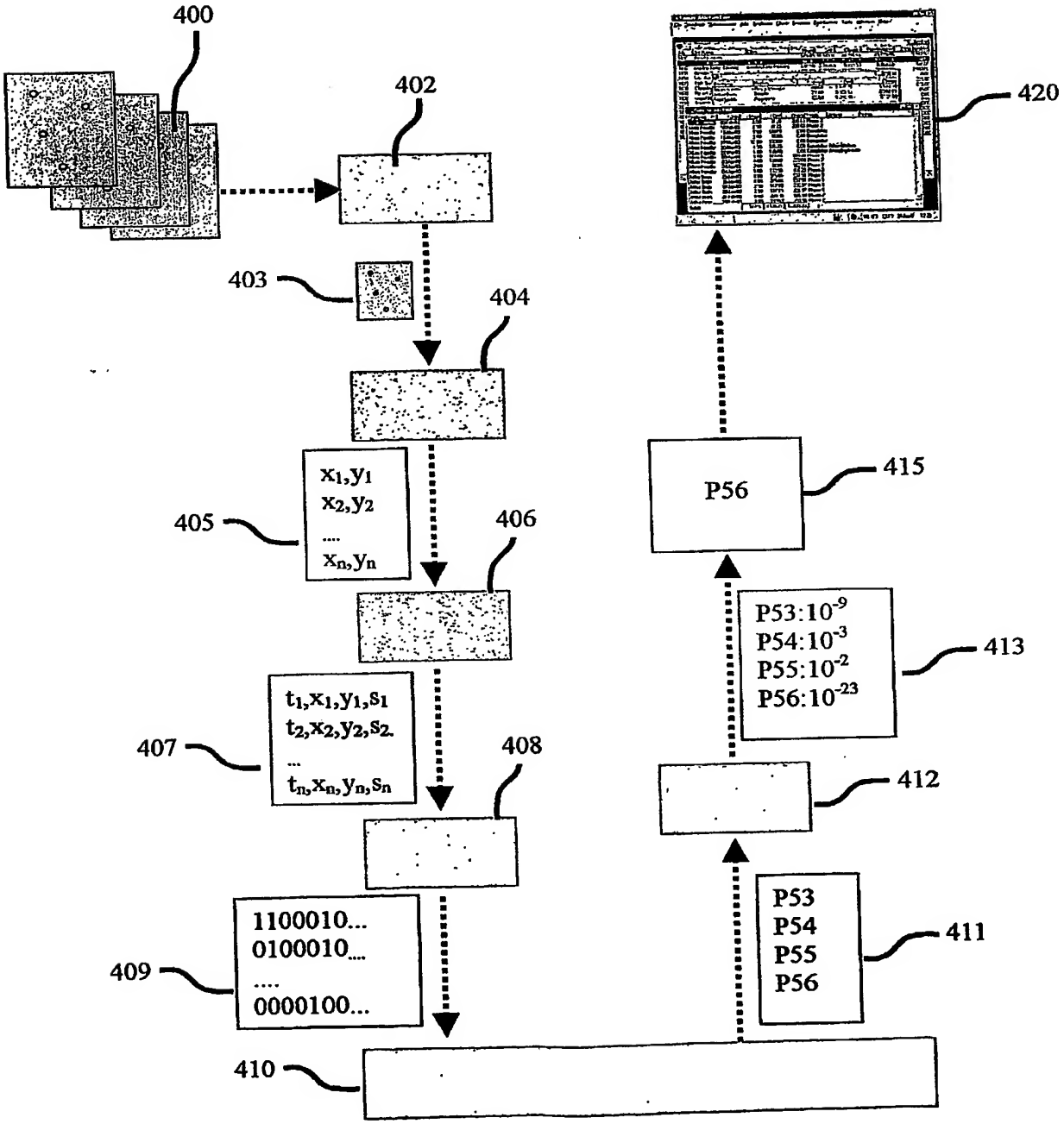

400
402
420
403
404
x_1,y_1
x_2,y_2
....
x_n,y_n
405
406
P56
415
t_1,x_1,y_1,s_1
t_2,x_2,y_2,s_2
...
t_n,x_n,y_n,s_n
407
408
P53:10^{-9}
P54:10^{-3}
P55:10^{-2}
P56:10^{-23}
413
412
1100010...
0100010....
....
0000100...
409
P53
P54
P55
P56
411
410

FIGURE 11